GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 14:17:16 ; Search time 303 Seconds
               (without alignments)
               9197.419 Million cell updates/sec

Title: US-10-060-066-2
Perfect score: 656
Sequence: 1 tatggcttcacctcgcacct...........tgtcatttttattctcttg 656

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:* .
               1: geneseqn1980s:*
               2: geneseqn1990s:*
               3: geneseqn0008:*
               4: geneseqn2001as:*
               5: geneseqn2001bs:*
               6: geneseqn2002s:*
               7: geneseqn2003as:*
               8: geneseqn2003bs:*
               9: geneseqn2003cs:*
               10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	96.4	14.7	1173	7	ABX72170	Abx72170 Human NOV
2	86.4	13.2	984	3	AA299791	Aa299791 DNA encod
3	86.4	13.2	984	3	AAF81301	Aaf81301 Human cac
4	83.2	12.7	1071	6	ABK51915	Abk51915 cDNA enco
5	73.2	11.2	1093	6	ABQ17109	Abq17109 Oligonucl
6	73.2	11.2	1093	6	ABQ17108	Abq17108 Oligonucl
7	54.4	8.3	972	3	AA299795	Aa299795 DNA encod
8	54.4	8.3	1252	4	ABA00049	Aba00049 Human vol
9	54.4	8.3	1478	5	AA251628	Aa251628 Human mem
10	53.6	8.2	1477	5	ABK51914	Abk51914 cDNA enco
11	52.8	8.0	1558	5	ABK51913	Abk51913 cDNA enco
12	52	7.9	945	3	AA299794	Aa299794 EST corre
13	52	7.9	948	3	AA299793	Aa299793 DNA encod
14	52	7.9	1383	4	ABA09004	Aba09004 Human vol
15	52	7.9	1854	6	AAC77216	Aac77216 Human ORF
16	52	7.9	3914	6	ABV76746	Abv76746 Human ple
17	45.8	7.0	110000	4	AAI95682_16	Aai95682_16
18	45.8	7.0	110000	4	AAI95683_16	Aai95683_16
19	43.6	6.6	451	6	ABT10409	Abt10409 Human bre
20	43	6.6	1031	9	ADD29576	Add29576 Human tum
21	42.8	6.5	799	2	AAV55831	Aav55831 Nucleotid
22	42.8	6.5	1925	2	AAX90924	Aax90924 Epstein B
23	42.8	6.5	1926	3	AAA50254	Aaa50254 Epstein B
24	42.8	6.5	1926	4	AAF82902	Aaf82902 EBV tethe
25	42.8	6.5	2580	6	AAA75454	Aaa75454 Nucleotid
26	42.8	6.5	2580	6	AAI64275	Aai64275 Epstein-B
27	42.8	6.5	5452	2	AAX90923	Aax90923 Anti-sens
28	42.8	6.5	8705	2	AAZ23778	Aaz23778 Vector pS
29	42.8	6.5	9600	2	AAV21683	Aav21683 Vector pl
30	42.8	6.5	10285	6	ABS71027	Abs71027 pCEP-Xa-F
31	42.8	6.5	10285	6	ABS66453	Abs66453 Plasmid p
32	42.8	6.5	10380	2	AAZ22248	Aaz22248 Nucleotid
33	42.8	6.5	10596	2	AAQ51731	Aaq51731 Plasmid p
34	42.8	6.5	10596	2	AAX15650	Aax15650 Nucleotid
35	42.8	6.5	10596	3	AAT40348	Aat40348 Plasmid p
36	42.8	6.5	16080	3	AAA59553	Aaa59553 DNA clone
37	42.8	6.5	585	6	ABQ50506	Abq50506 Oligonucl
38	42.6	6.5	585	6	ABQ50507	Abq50507 Oligonucl
39	42.6	6.5	1000	3	AAA02484	Aaa02484 Human col
40	42.4	6.4	8837	6	ADD25495	Add25495 Binding d
41	42.2	6.4	615	7	ACA23975	Aca23975 Prokaryot
42	41.8	6.4	415	4	AAI82119	Aai82119 Human pol
43	41.6	6.3	541	1	AAC07568	Aac07568 Human sec
44	41.6	6.3	1646	3	AAZ35999	Aaz35999 S. kitasa
45	41.4	6.3	12733	6	ABK98631	Abk98631 Vector pE

ALIGNMENTS

RESULT 1
ABX72170
ID   ABX72170 standard; cDNA; 1173 BP.
XX
AC   ABX72170;
XX
DT   03-JUN-2003 (first entry)
XX
DE   Human NOVX polynucleotide #1.
XX
KW   Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW   hypertension; congenital heart defect; aortic stenosis; valve disease;
KW   atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW   pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW   tuberous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW   obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease; immune disorder; haematopoietic disorder;
KW   haemophilia; hypercoagulation; Crohn's disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200281498-A2.
XX
PD   17-OCT-2002.
XX
PF   03-APR-2002; 2002WO-US010780.
XX
PR   03-APR-2001; 2001US-0281086P.
PR   03-APR-2001; 2001US-0281136P.
PR   05-APR-2001; 2001US-0281863P.
PR   05-APR-2001; 2001US-0281906P.
PR   06-APR-2001; 2001US-0282020P.
PR   10-APR-2001; 2001US-0282030P.
PR   10-APR-2001; 2001US-0282934P.
PR   12-APR-2001; 2001US-0283512P.
PR   13-APR-2001; 2001US-0283710P.
PR   17-APR-2001; 2001US-0284234P.
PR   19-APR-2001; 2001US-0285325P.
PR   20-APR-2001; 2001US-0285381P.
PR   20-APR-2001; 2001US-0285590P.
PR   23-APR-2001; 2001US-0285748P.
PR   23-APR-2001; 2001US-0285890P.
PR   24-APR-2001; 2001US-0286086P.
PR   25-APR-2001; 2001US-0286292P.
PR   25-APR-2001; 2001US-0287130P.
PR   27-APR-2001; 2001US-0287130P.
PR   02-MAY-2001; 2001US-0288257P.

Left column

```
PR   29-MAY-2001;  2001US-0294164P.
PR   30-MAY-2001;  2001US-0294484P.
PR   18-JUN-2001;  2001US-0298952P.
PR   19-JUN-2001;  2001US-0299237P.
PR   19-JUN-2001;  2001US-0299276P.
PR   12-SEP-2001;  2001US-0318750P.
PR   25-SEP-2001;  2001US-0324800P.
PR   25-SEP-2001;  2001US-0324802P.
PR   27-SEP-2001;  2001US-0325684P.
PR   17-OCT-2001;  2001US-0330143P.
PR   14-NOV-2001;  2001US-0332131P.
PR   14-NOV-2001;  2001US-0332240P.
PR   21-NOV-2001;  2001US-0332799P.
PR   21-NOV-2001;  2001US-0332115P.
PR   04-DEC-2001;  2001US-0337621P.
PR   03-JAN-2002;  2002US-0345783P.
PR   16-JAN-2002;  2002US-0350251P.
PR   02-APR-2002;  2002US-0011427O.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
PI   Patturajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD;
PI   Gorman L, Shenoy SG, Pena CEA, Smithson G, Burgess CE, Gerlach V;
PI   Padigaru M, Shimkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;
PI   Anderson DW, Lette MW, Rastelli L, Edinger SR, Stone DJ, Ji W;
PI   Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
PI   Ellerman K;
XX
PS   Claim 17; Page 91; 66pp; English.
XX
PT   New isolated NOVX polypeptide useful for treating atherosclerosis,
PT   metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT   neurodegenerative disorders, Alzheimer's disease and cancer.
XX
DR   WPI; 2003-046858/04.
DR   P-PSDB; ABU54542.
XX
CC   The invention relates to human polypeptides, termed NOVX, and the
CC   polynucleotides encoding them. The polypeptides and polynucleotides are
CC   useful for diagnosing disease, and screening for potential therapeutic
CC   agents. The sequences are useful for treating metabolic disorders,
CC   cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC   stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC   ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC   septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma,
CC   atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC   disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC   haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC   and cancer. Sequences ABX72170-ABX72275 represent human NOVX
CC   polynucleotides of the invention
XX
SQ   Sequence 1173 BP; 246 A; 386 C; 347 G; 194 T; 0 U; 0 Other;

Query Match             14.7%;  Score 96.4;  DB 7;  Length 1173;
Best Local Similarity   83.8%;
Matches  109;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         187 TGAAACGGCTGGGAATGAGAGAGGGGGGTTTGTTGGTGTGAAAAGGGCCGTTCAGGTTACTACTGA 246
               |||||  ||||||||   |||| || |  ||||| |||  |||||||||||||||||||||||||||
Db          38 TGAAGGCGGTGGAAACGAAAATGCGGGCGGGGCCCTCTCGGTGCGAGAAGGGGGTGCAGGTGCTGCTGA 97

Qy         247 CCACCATCGGGCGCGCCCTTCGCGGCGTTTGGCCTCATGACCATCGCCATCAGCAGCACTGACTACT 306
               | |||||||||||| |||| |||||  |||||||||||||||||||||||||||||  |||||| |||
Db          98 CGACGGTGGGCGGCCCTTCCCGCGCCTTCGCGCCTCATGACCATCGCCATCAGCACTGACTPACT 157

Qy         307 GGCTCTACAC 316
               ||||||||||
Db         158 GGCTCTACAC 167
```

Right column

```
ID   AAZ99791 standard; DNA; 984 BP.
XX
AC   AAZ99791;
XX
DT   12-JUL-2000  (first entry)
XX
DE   DNA encoding a voltage-gated calcium channel designated CACNGLIKE3.
XX
KW   Human; skeletal muscle; voltage-gated calcium channel; CACNGLIKE3;
KW   neurological disorder; epilepsy; stroke; head trauma; migraine;
KW   affective disorder; depression; anxiety; schizophrenia; pain; cancer;
KW   neurodegenerative disorder; Alzheimer's disease; cognitive disorder;
KW   chromosome localisation; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..984
FT                   /*tag= a
FT                   /product= "voltage-gated calcium channel"
XX
PN   WO200014224-A1.
XX
PD   16-MAR-2000.
XX
PF   06-SEP-1999;  99WO-GB002944.
XX
PR   08-SEP-1998;  98GB-00019592.
XX
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Duckworth DM, Hayes PD;
XX
DR   WPI; 2000-256976/22.
DR   P-PSDB; AAY84372.
XX
PS   Claim 4; Page 31; 37pp; English.
XX
PT   Isolated voltage-gated calcium channel polypeptide, designated
PT   CACNGLIKE3, to treat diseases such as neurological disorders, epilepsy,
PT   neurodegenerative disorders, cognitive disorders and cancer; comprises
PT   327 amino acid sequence.
XX
CC   The present sequence encodes a human skeletal muscle voltage-gated
CC   calcium channel polypeptide, designated CACNGLIKE3. The CACNGLIKE3
CC   polypeptide and polynucleotide are useful in the treatment of diseases
CC   such as neurological disorders, epilepsy, stroke, head trauma, migraine,
CC   affective disorders including depression and anxiety, schizophrenia,
CC   neurodegenerative disorders including Alzheimer's disease, cognitive
CC   disorders, types of pain and cancer. The polynucleotide is also valuable
CC   for chromosome localisation studies. The CACNGLIKE3 polypeptide and
CC   polynucleotide are also useful in diagnostic assays for detecting
CC   diseases associated with inappropriate CACNGLIKE3 activity or levels
XX
SQ   Sequence 984 BP; 206 A; 319 C; 265 G; 194 T; 0 U; 0 Other;

Query Match             13.2%;  Score 86.4;  DB 3;  Length 984;
Best Local Similarity   66.8%;
Matches  123;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy         214 TGTGGTGTGAAAAGGGCCGTTCAGGTACTACTACTGACCACCATCGGGCGCGCCCTTCGCTTTG 273
               || | ||| |||| ||||| |||  |||||||||||||||||||||||||||||| ||||||| |
Db           5 TGCGATGGGAGCCGCGGCGCTGCAGATGCTGGGTGCTGACCACGCGCCGGAGCCTTCGCGCCTTCT 64

Qy         274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACAACAAGAGCTCTCATCTCGA 333
               |||||||| ||||||||| |||| |||||| ||||| ||| |  |||||    |||| ||||
Db          65 CGCTCATGGCCATCGCCACCAGCTACTTGGCGTTACTCGACGACCATCTGA 124

Qy         334 ACACCACCCAACCTCACACGACAGGTGATGACGGACCACCCCCATCGTGGGCGCCAGTGGGCTCCT 393
               || |||| |||||| ||| ||| |||   |||||   |||||||  || ||||| ||    ||| |
Db         125 ACGGCACCAACCTGACCATGGACGACGGGCCCCGCCCCGCGGCGGGGGCGGCGACGACTTCA 184

Qy         394 CCGA 397
               ||||
Db         394 CCGA 397
```

4

```
Db
                    185 CCCA 188
                        |||
Qy                  185 CCCA 188

Db

RESULT 3
AAF81301
ID   AAF81301 standard; cDNA; 984 BP.
XX
AC   AAF81301;
XX
DT   01-JUN-2001  (first entry)
XX
DE   Human Cacng8 cDNA.
XX
KW   Human; stargazin-like; CACNG8; anticonvulsant; hypotensive; vasotropic;
KW   antiarrhythmic; antianginal; analgesic; nootropic; cytostatic;
KW   neuroprotective; cerebroprotective; antidiabetic; antimigraine;
KW   antiarteriosclerotic; immunosuppressive; antiparkinsonian; auditory;
KW   ophthalmological; tranquiliser; neuronal disorder;
KW   calcium channel modulation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200121791-A2.
XX
PD   29-MAR-2001.
XX
PF   25-SEP-2000; 2000WO-GB003685.
XX
PR   23-SEP-1999; 99GB-0022571.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Clare JJ, Plumpton M, Moss FJ, Sanseau P;
XX
DR   WPI; 2001-266074/27.
DR   P-PSDB; AAB73979.
XX
PT   Novel stargazin-like polypeptides capable of modulating the steady state
PT   inactivation of an alpha1 pore-forming subunit of a voltage-gated calcium
PT   channel, for treating central nervous system disorders.
XX
PS   Claim 6; Page 38-39; 48pp; English.
XX
CC   The present sequence is provided in a specification relating to an
CC   isolated stargazin-like polypeptide comprising a sequence of 327 or 275
CC   amino acids fully defined in the specification, or its variant capable of
CC   modulating the steady state. The polypeptide is useful for treating a
CC   subject having a neuronal disorder that is responsive to calcium channel
CC   modulation. It is useful in the manufacture of a medicament for treatment
CC   or prophylaxis of a disorder that is responsive to calcium channel
CC   modulation, e.g. epilepsy, episodic ataxia, spinocerebellar ataxia,
CC   hypertension, ischaemic heart disease, arrhythmia, angina, pain, cerebral
CC   ischaemia, Alzheimer's disease, neuroprotection, stroke, diabetes,
CC   cerebral vasospasm, atherosclerosis, tardive dyskinesias, peripheral
CC   vascular disease, immunosuppression, cancerous diseases, migraine,
CC   headache, bipolar disorder, unipolar depression, anxiety, Parkinson's
CC   disease, cognitive disorders, ophthalmic diseases, neuromuscular
CC   disorders and tinnitus. Nucleic acids encoding these polypeptides are
CC   useful in the production of the polypeptides, and as primers.
XX
SQ   Sequence 984 BP; 204 A; 319 C; 265 G; 194 T; 0 U; 2 Other;

Query Match             13.2%;  Score 86.4;  DB 4;  Length 984;
Best Local Similarity   66.8%;                Pred. No.1.5e-12;
Matches   123;  Conservative   0;  Mismatches  61;  Indels  0;  Gaps  0;

Qy   214  TGTGGTGTGAAAAGGGCGTTCAGGTACTACTACGGCCATCGGGCCTTCGCGGGCTTTTG  273
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   5    TGGCAATGGCACCCGGCGGGCGTGCACAGATGCTGCTGACCACGCGCCGGAGCCTTCT  64

Qy   274  GCCTCCATGACCATCCAGCACTGACTACTGGCTTCTCACACAAGAGAGCTCTCATCTGCA  333
          |||||||||||||||||||||||||||||||||||||||||||||||
```

```
CC   65   CGCTCCATGGCCATCCCCATGGGCACCGACTTACTGCTGTTACTCCAGCGCCACACTGCA  124
          ||||||||||||||||||||||||||||||||||| |||||| |||||||||| |||||
Db   334  ACACCACCACCTCACAGCAGGTGATGACGGACGAGGATCAGCGCCACCACCCCATCGTTG  393

Db   125  ACGGCACCAACCTGACCATGGACGACGGGCCCCCCGCCCCCGCGCCCGCCCGGCGCCCGC  184
          |||||||||||||||||||||||||||||||||||||||||||||||
Qy   394  CCGA  397
Db   185  CCCA  188

RESULT 4
ABKS1915
ID   ABKS1915 standard; cDNA; 1071 BP.
XX
AC   ABKS1915;
XX
DT   13-AUG-2002  (first entry)
XX
DE   cDNA encoding mouse Cacng4.
XX
KW   Mouse; neuronal voltage-gated calcium channel gamma subunit;
KW   autoimmune disease; Lambert-Eaton syndrome; neuronal disease; epilepsy;
KW   immunosuppressive; anticonvulsant; Cacng4; gene; ss.
XX
OS   Murinae gen. sp.
XX
FH   Key             Location/Qualifiers
FH
FT   5'UTR           1..21
FT                   /*tag= a
FT   CDS             22..1005
FT                   /product= "Cacng4 #1"
FT                   /*tag= b
FT   CDS             22..1005
FT                   /product= "Cacng4 #2"
FT                   /*tag= c
FT                   /transl_except=(pos:979..981, aa:Xaa)
FT                   /transl_except=(pos:991..993, aa:Xaa)
FT                   /note="Xaa= unknown"
FT   3'UTR           1006..1071
FT                   /*tag= d
XX
PN   US6365337-B1.
XX
PD   02-APR-2002.
XX
PF   27-JUL-1998; 98US-00123030.
XX
PR   27-JUL-1998; 98US-00123030.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
PA   (JACK-) JACKSON LAB.
XX
PI   Letts VA, Frankel WN, Campbell KP, Felix R, Biddlecome G;
XX
DR   WPI; 2002-433421/46.
DR   P-PSDB; AAU97155, AAU97156.
XX
PT   Novel nucleic acid sequences encoding a neuronal voltage-gated calcium
PT   channel gamma subunit useful in screening for compounds which modulate
PT   activity of the channel and in diagnosing, treating neuronal diseases.
XX
PS   Disclosure; Fig 7; 36pp; English.
XX
CC   The present invention relates to the isolation of mouse genes encoding
CC   neuronal voltage-gated calcium channel gamma subunits designated Cacng2,
CC   Cacng3, and Cacng4. The genes are useful for identifying candidate
CC   compounds for modulating the activity of human neuronal voltage-gated
CC   calcium channels. They are also useful for diagnosing and treating the
CC   autoimmune disease Lambert-Eaton syndrome, as well as diagnosing defects
CC   in gamma subunit genes of a patient with a neuronal disease such as
CC   epilepsy. The present sequence encodes mouse Cacng4. Note: Nucleotides 1-
CC   21 of the present sequence are not shown in Fig 7 but are included in the
```

```
CC   sequence shown in the sequence listing
XX
SQ   Sequence 1071 BP; 227 A; 345 C; 284 G; 213 T; 0 U; 2 Other;
     Query Match             12.7%;  Score 83.2;  DB 6;  Length 1071;
     Best Local Similarity   65.8%;  Pred. No. 1.1e-11;
     Matches  121;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy        214 TGTGGGTGTGAAAGGCGTTCAGGTACTACTACGACCACCATCGGCGCCGCTTGCGGGCTTTG 273
                                    |||
Db         26 TCGGATCGGACCCCGGCGTGCAGATGCTGCTGACCACCATCGGGCGCCCAGCCCTTCT 85

Qy        274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCGTTCTCACAAGAGCTCTCATCTGCA 333
                    ||||     ||     |||||||        ||||||||     ||||||||   |||
Db         86 CGCTCATGGCCATCGACCATCGGGCTGTTACTTCCAGCGCGCCGCCACATCTGCA 145

Qy        334 ACACCACCAACCTTCACAGCAGGTTGATGACGGACCACCCCCATTCGTGGGGGGCCAGTGGCTCCT 393
                  ||||||    |||  |||     ||||||     ||||||||     ||||||     |||
Db        146 ACGGCACCAAACCTTGACCATTGACGACGACGGGCCCCCCCGCGCGCTCGGCGCGACCTCA 205

Qy        394 CCGA 397
                ||
Db        206 CCCA 209

RESULT 5
ID   ABQ17109 standard; DNA; 1093 BP.
XX
AC   ABQ17109;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 3700.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:

CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP;/8); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 1093 BP; 439 A; 406 C; 153 G; 95 T; 0 U; 0 Other;
     Query Match             11.2%;  Score 73.2;  DB 6;  Length 1093;
     Best Local Similarity   65.1%;  Pred. No. 4.4e-09;
     Matches  108;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy        232 TTCAGGTNACTACTGACCACCACCATCGGCGCCGCTTCGCGGGCGCTTTTGGCCTCATGACCATCGCCA 291
                     |||||||     ||||||||     |||||||||     |||||     |||||||||
Db        411 TACAAAATACTACTNAACCACGACCGAAAACTTCGCCGCCGCTTCGCGGGCTTCATNAACCATCGCCA 470

Qy        292 TCAGCACTGACTACTGGCTCTCACACAAGAGACGCTCTCATTCGTGCAACACACCAACCTCACAG 351
                    ||||||||||     |||        ||||||     |||||     |||||     |||
Db        471 TCGACACGACTACTAACTATACTCCAACGCGCACACAATCTACAGACGACCAACCTTACCA 530

Qy        352 CAGGTGATGACGGACCACCACCCCCATTCGTGGGGGGCCAGTGGGCTTCCTCCGA 397
                    |||||||     |||||||     ||||||     |||||     ||||
Db        531 TAAACGACGAACCCCGCCGCCCGCGCCGCGCCGCCGCGACGACCTCACCA 576

RESULT 6
ABQ17108/c
ID   ABQ17108 standard; DNA; 1093 BP.
XX
AC   ABQ17108;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 3699.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
```

CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 1093 BP; 95 A; 153 C; 406 G; 439 T; 0 U; 0 Other;

Query Match   11.2%; Score 73.2; DB 6; Length 1093;
Best Local Similarity   65.1%;
Matches 108; Conservative   0; Mismatches 58; Indels   0; Gaps   0;

Qy   232 TTCAGGTACTACTGACCACCACCATCCGGCGGGGGCCTTCCGGGGCCTTTTGGCCTCATGACCATCGGCCA 291
Db   683 TACAAATACTACTAACCACGACCCAAACCTTCGCGCGCCCTTCTGCCTCATRAACCATCGGCCA 624

Qy   292 TCAGCACTGACTACTGGCCTCTCACACAGAGCTCTCATCTGCAACACCACCAACCTCACAG 351
Db   623 TCGACACCGACTACTAACTATATACTCCAACGCGCACATCTACACGACGACCACCAACCTAACCA 564

Qy   352 CAGGTGATGACGGACCACCACCCATCGTGTGGGGGGCCACTGGCCTCCTCCGA 397
Db   563 TAAACGACGAACCCCGCCCGCGCCCGGCCGGCCGGGGACGACCTCACCCA 518

RESULT 7
ID AA299795 standard; DNA; 972 BP.
XX
AC AA299795;
XX
DT 12-JUL-2000 (first entry)
XX
DE DNA encoding a voltage-gated calcium channel designated CACNGLIKE1.
XX
KW Human; voltage-gated calcium channel; CACNGLIKE1; neurological disorder;
KW epilepsy; stroke; head trauma; migraine; affective disorder; depression;
KW anxiety; schizophrenia; neurodegenerative disorder; Alzheimer's disease;
KW cognitive disorder; pain; cancer; chromosome localization; ss.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   1..972
FT   /tag= a
FT   /product= "voltage-gated calcium channel"
XX
PN WO200014223-A1.
XX
PD 16-MAR-2000.
XX
PF 06-SEP-1999; 99WO-GB002937.
XX
PR 07-SEP-1998; 98GB-00019474.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Duckworth DM, Hayes PD;
XX
DR WPI; 2000-256975/22.
DR P-PSDB; AAY84376.
XX
PT New human voltage-gated calcium channel, known as CACNGLIKE1, useful in

---

PT the treatment of diseases such as neurological disorders, epilepsy,
PT neurodegenerative disorders, cognitive disorders and cancer, comprises a
PT 323 amino acid sequence.
XX
PS Claim 4; Page 30; 35pp; English.
XX
CC The present sequence encodes a human voltage-gated calcium channel, known
CC as CACNGLIKE1. The CACNGLIKE1 polypeptide and polynucleotide are useful
CC in the treatment of diseases such as neurological disorders, epilepsy,
CC stroke, head trauma, migraine, affective disorders including depression
CC and anxiety, schizophrenia, neurodegenerative disorders including
CC Alzheimer's disease, cognitive disorders, types of pain and cancer. The
CC CACNGLIKE1 polynucleotide is also valuable for chromosome localization
CC studies. CACNGLIKE1 polypeptide and polynucleotide are also useful in diagnostic
CC assays for detecting diseases associated with inappropriate CACNGLIKE1
CC activity or levels
XX
SQ Sequence 972 BP; 231 A; 311 C; 231 G; 199 T; 0 U; 0 Other;

Query Match   8.3%; Score 54.4; DB 3; Length 972;
Best Local Similarity   70.2%;
Matches   73; Conservative   0; Mismatches 31; Indels   0; Gaps   0;

Qy   218 GTGTGAAAAGGGCGTTCAGGTACTACTGACCACCATCGGGCCGCGCGGCGGCGTTTTGGCCT 277
Db   9 GTTTGATCGAGGGTGTTCAAATGCTTTTAACCACCGTTGGGTGCTTCGGCTGCCTTCAGCCT 68

Qy   278 CATGACCATCGCCATCAGCACGACTGACTACTCGGCTCTACACAGAGAG 321
Db   69 GATGACCATAGCCTGTGGGAACCGACTATTGGCTCTACTCCAGAG 112

RESULT 8
ID ABA09049 standard; cDNA; 1252 BP.
XX
AC ABA09049;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human voltage gated Ca channel subunit homologue cDNA, SEQ ID NO:825.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR P-PSDB; ABB11805.

```
XX
PT   Human proteins and DNA encoding sequences useful for preventing, treating
PT   or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT   and cancer.
XX
PS   Claim 1; Page 731-732; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
CC   nucleotide of the invention, methods of producing the novel polypeptides,
CC   antibodies against the polypeptides, and methods of detecting the nucleotides
CC   or polypeptides in a sample, and methods of identifying compounds which
CC   bind to polypeptides of the invention. Although novel, many of the
CC   polypeptides of the invention have homology to known proteins, thereby
CC   giving an insight into their probable biological activities, and hence
CC   potential therapeutic applications. The polypeptides of the invention may
CC   have various activities, including cytokine, cell growth factor or cell
CC   differentiation activities; stem cell growth factor activity;
CC   haematopoiesis regulatory activity; tissue growth activity;
CC   immunomodulatory activity; activin- or inhibin-related activities;
CC   chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC   thrombolytic activities; receptor or ligand activities; or may be
CC   involved in oncogenesis, cancer cell proliferation or metastasis.
CC   Depending on their biological activities, polypeptides and nucleotides of
CC   the invention are useful for preventing, treating or ameliorating medical
CC   conditions, e.g. by protein or gene therapy. Such conditions include
CC   cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
CC   disorders), chronic inflammatory conditions (e.g. asthma or arthritis),
CC   proliferative retinopathy, atherosclerosis, coronary heart disease,
CC   arterial ischaemia, bone disorders (e.g. osteoporosis), and abnormal
CC   vascular growth. Polypeptides involved with tissue regeneration and
CC   repair (or nucleic acids encoding them) may be used to promote wound
CC   healing (e.g. of burns, incisions and ulcers), while those with
CC   immunomodulatory activities may be used in the treatment of viral,
CC   bacterial and fungal infections in addition to immune disorders.
CC   Polypeptides with growth factor activity may be used in cell cultures to
CC   promote cell growth. For example, such polypeptides may be used to
CC   manipulate stem cells in culture to give rise to neuroepithelial cells
CC   that can be used to augment or replace cells damaged by illness,
CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a cDNA encoding a
CC   novel human polypeptide of the invention
XX
SQ   Sequence 1252 BP; 302 A; 371 C; 266 G; 313 T; 0 U; 0 Other;

Query Match             8.3%;  Score 54.4;  DB 4;  Length 1252;
Best Local Similarity   70.2%;
Matches   73;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy        218 GTGTGAAAAGGCGGTTCAGGTAGTACTACTGACCACCACCATCGGCGCGGCGGCCTTCGCGGCGGCCTTTTGGCCT 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||  ||||  ||||||
Db        291 GTTTGATCGAGGTGTTCAAAATGCTTTTAACCACCGTTGGTGCTTTCGCTGCGGCCTTCAGCCT 350

Qy        278 CATGACCATCGCCATCAGCACTACTACTGGGCTCTACACAAGAG 321
              ||||||| |||||||| ||||||||||||||||||| |||
Db        351 GATGACCATAGCTGTGGGAACCGACTATTGGCTCTACTCCAGAG 394

RESULT 9
AAZ51628
ID   AAZ51628 standard; cDNA; 1478 BP.
XX
AC   AAZ51628;
XX
DT   21-JUN-2000  (first entry)
XX
DE   Human membrane channel protein-12 (MECHP-12) cDNA.
XX
KW   Membrane channel protein-12; MECHP-12; diagnosis; treatment; lymphoma;
KW   cell proliferative disorder; bursitis; atherosclerosis; cancer; sarcoma;
KW   inflammatory disorder; AIDS; Addison's disease; cystic fibrosis; asthma;
KW   diabetes mellitus; osmoregulatory disorder; diarrhoea; renal failure;
KW   muscular disorder; myocarditis; Duchenne's muscular dystrophy; nootropic;
KW   cardiovascular disorder; hypertension; bronchitis; vasculitis; cardiant;
KW   neurological disorder; Alzheimer's disease; Parkinson's disease; human;
KW   Huntington's disease; antiarteriosclerotic; hepatotropic; cytostatic;
KW   anti-HIV; antianaemic; neuroprotective; immunomodulator; antidiabetic;
KW   hypotensive; vasotropic; antiasthmatic; antiinflammatory; antidepressant;
KW   anticonvulsant; thrombolytic; antiParkinsonian; immunostimulant; as.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             364..1335
FT                   /*tag= a
FT                   /product= "MECHP-12"
FT                   994..1041
FT                   /*tag= b
FT   misc_binding    /bound_moiety= "Primer or Probe"
XX
PN   WO200012711-A2.
XX
PD   09-MAR-2000.
XX
PF   02-SEP-1999; 99WO-US020468.
XX
PR   02-SEP-1998; 98US-0155226P.
PR   12-NOV-1998; 98US-00191283.
PR   09-DEC-1998; 98US-0155255P.
PR   26-JAN-1999; 99US-0155211P.
PR   10-FEB-1999; 99US-0155263P.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Au-Young J,  Bandman O,  Tang YT,  Reddy R,  Hillman JL,  Yue H;
PI   Lal P,  Corley NC,  Guegler KJ,  Gorgone G,  Baughn MR,  Azimzai Y;
XX
DR   WPI; 2000-256643/22.
DR   P-PSDB; AAY70462.
XX
PT   Novel human membrane channel protein and polynucleotide useful for
PT   diagnosing and treating cell proliferative, inflammatory, secretory,
PT   osmoregulatory, muscular, cardiovascular and neurological disorders.
XX
PS   Claim 9; Page 126; 140pp; English.
XX
CC   The present sequence is a cDNA identified in Incyte clone 1267774 derived
CC   from BRAINOT09 cDNA library. It encodes human membrane channel protein-12
CC   (MECHP-12), which is expressed in nervous tissues. Anti-MECHP antibodies
CC   can be used as therapeutic antagonists and reagents for diagnosis and
CC   monitoring diseases. MECHP cDNA can be used for diagnosis of MECHP-
CC   related diseases and gene mapping. MECHP can be used for treatment of
CC   cell proliferative disorders such as bursitis and atherosclerosis,
CC   cancers like lymphoma and sarcoma, inflammatory disorders like AIDS and
CC   Addison's disease, transport/secretory disorders like cystic fibrosis and
CC   diabetes mellitus, osmoregulatory disorders like diarrhoea and renal
CC   failure, muscular disorders like myocarditis and Duchenne's muscular
CC   dystrophy, cardiovascular disorders like hypertension and vasculitis,
CC   congenital lung anomalies like bronchitis and asthma and neurological
CC   disorders like Alzheimer's disease, Parkinson's disease and Huntington's
CC   disease
XX
SQ   Sequence 1478 BP; 374 A; 416 C; 339 G; 349 T; 0 U; 0 Other;

Query Match             8.3%;  Score 54.4;  DB 3;  Length 1478;
Best Local Similarity   70.2%;
Matches   73;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy        218 GTGTGAAAAGGCGGTTCAGGTAGTACTACTGACCACCACCATCGGCGCGGCGGCCTTCGCGGCGGCCTTTTGGCCT 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||  ||||  ||||||
Db        372 GTTTGATCGAGGTGTTCAAAATGCTTTTAACCACCGTTGGTGCTTTCGCTGCGGCCTTCAGCCT 431

Qy        278 CATGACCATCGCCATCAGCACTACTACTGGGCTCTACACAAGAG 321
              ||||||| |||||||| ||||||||||||||||||| |||
Db            GATGACCAT
```

Db   432   GATGACCATAGCTGTGGGAACCGACTATTGGCTCTACTCCAGAG   475

RESULT 10
ABK51914
ID   ABK51914 standard; cDNA; 1477 BP.
XX
AC   ABK51914;
XX
DT   13-AUG-2002 (first entry)
XX
DE   cDNA encoding mouse Cacng3.
XX
KW   Mouse; neuronal voltage-gated calcium channel gamma subunit;
KW   autoimmune disease; Lambert-Eaton syndrome; neuronal disease; epilepsy;
KW   immunosuppressive; anticonvulsant; Cacng3; gene; ss.
XX
OS   Murinae gen. sp.
XX
FH   Key   Location/Qualifiers
FH
FT   5'UTR   1..487
FT   /*tag= a
FT   CDS   488..1435
FT   /*tag= b
FT   /product= "Cacng3"
FT   3'UTR   1436..1477
FT   /*tag= c
XX
PN   US6365337-B1.
XX
PD   02-APR-2002.
XX
PF   27-JUL-1998;   98US-00123030.
XX
PR   27-JUL-1998;   98US-00123030.
XX
PA   (IOWA) UNIV IOWA RES FOUND.
PA   (JACK-) JACKSON LAB.
XX
PI   Letts VA,   Frankel WN,   Campbell KP,   Felix R,   Biddlecome G;
XX
DR   WPI; 2002-433421/46.
DR   P-PSDB; AAU97154.
XX
PT   Novel nucleic acid sequences encoding a neuronal voltage-gated calcium
PT   channel gamma subunit useful in screening for compounds which modulate
PT   activity of the channel and in diagnosing, treating neuronal diseases.
XX
PS   Disclosure; Fig 6; 36pp; English.
XX
CC   The present invention relates to the isolation of mouse genes encoding
CC   neuronal voltage-gated calcium channel gamma subunits designated Cacng2,
CC   Cacng3, and Cacng4. The genes are useful for identifying candidate
CC   compounds for modulating the activity of human neuronal voltage-gated
CC   calcium channels. They are also useful for diagnosing and treating the
CC   autoimmune disease Lambert-Eaton syndrome, as well as diagnosing defects
CC   in gamma subunit genes of a patient with a neuronal disease such as
CC   epilepsy. The present sequence encodes mouse Cacng3
XX
SQ   Sequence 1477 BP; 352 A; 449 C; 356 G; 320 T; 0 U; 0 Other;

Query Match   8.2%;   Score 53.6; DB 6; Length 1477;
Best Local Similarity   66.4%;
Matches   77;   Conservative   0;   Mismatches   39;   Indels   0;   Gaps   0;

Qy   218   GTGTGAAAAGGGCCGTTCAGTACTACTGACCACCATCGGCGCGCCCTTCGGCGGCTTTTGGCCT   277
            ||||| ||| || ||||||| ||||| ||||||||||| |||||| |||||| |||
Db   496   GTGTGACAGAGGTATCCAGATGTTGATCACTACTGTAGGAGCCCTTCGCAGCTTTTAGTTT   555

Qy   278   CATGACCATCGCCATCACGACTACTGGCTCTACACAAGAGCTCTCATCTGCA   333
            ||||||||| ||||| ||||| |
Db   556   AATGACCATTGCACTGGGCGACGGGACTACTCGCTATATTCCAGAGGTGTGTGCACGA   611

Qy   ↓

RESULT 11
ABK51913
ID   ABK51913 standard; cDNA; 1558 BP.
XX
AC   ABK51913;
XX
DT   13-AUG-2002 (first entry)
XX
DE   cDNA encoding mouse Cacng2.
XX
KW   Mouse; neuronal voltage-gated calcium channel gamma subunit;
KW   autoimmune disease; Lambert-Eaton syndrome; neuronal disease; epilepsy;
KW   immunosuppressive; anticonvulsant; Cacng2; gene; ss.
XX
OS   Murinae gen. sp.
XX
FH   Key   Location/Qualifiers
FH
FT   5'UTR   1..389
FT   /*tag= a
FT   CDS   390..1361
FT   /*tag= b
FT   /product= "Cacng2"
FT   3'UTR   1362..1558
FT   /*tag= c
XX
PN   US6365337-B1.
XX
PD   02-APR-2002.
XX
PF   27-JUL-1998;   98US-00123030.
XX
PR   27-JUL-1998;   98US-00123030.
XX
PA   (IOWA) UNIV IOWA RES FOUND.
PA   (JACK-) JACKSON LAB.
XX
PI   Letts VA,   Frankel WN,   Campbell KP,   Felix R,   Biddlecome G;
XX
DR   WPI; 2002-433421/46.
DR   P-PSDB; AAU97153.
XX
PT   Novel nucleic acid sequences encoding a neuronal voltage-gated calcium
PT   channel gamma subunit useful in screening for compounds which modulate
PT   activity of the channel and in diagnosing, treating neuronal diseases.
XX
PS   Claim 2; Fig 2; 36pp; English.
XX
CC   The present invention relates to the isolation of mouse genes encoding
CC   neuronal voltage-gated calcium channel gamma subunits designated Cacng2,
CC   Cacng3, and Cacng4. The genes are useful for identifying candidate
CC   compounds for modulating the activity of human neuronal voltage-gated
CC   calcium channels. They are also useful for diagnosing and treating the
CC   autoimmune disease Lambert-Eaton syndrome, as well as diagnosing defects
CC   in gamma subunit genes of a patient with a neuronal disease such as
CC   epilepsy. The present sequence encodes mouse Cacng2
XX
SQ   Sequence 1558 BP; 418 A; 441 C; 360 G; 339 T; 0 U; 0 Other;

Query Match   8.0%;   Score 52.8; DB 6; Length 1558;
Best Local Similarity   69.2%;
Matches   72;   Conservative   0;   Mismatches   32;   Indels   0;   Gaps   0;

Qy   218   GTGTGAAAAGGGCCGTTCAGTACTACTGACCACCATCGGCGCGCCCTTCGGCGGCTTTTGGCCT   277
            |||| | || |||||||| |||| |||||||| ||||| |||||| |||||| |||
Db   398   GTTTGATCGAGGTGTTCAAATGCTTTTAACCACCGTTGGTGGTTTGCTGCCTTCAGCTT   457

Qy   278   CATGACCATCGCCATCACGACTACTGGCTCTACACAAGAG   321
            |||||||||| |||| ||| ||||| |
Db   458   GATGACCATGCTGTGGGAACCGACTATTGGCTGTACTCCAGAG   501

RESULT 12

```
ID   AA299794 standard; DNA; 945 BP.
XX
AC   AA299794;
XX
DT   12-JUL-2000  (first entry)
XX
DE   EST corresponding to a neuronal voltage-gated calcium chanel DNA.
XX
KW   Human; neuronal voltage-gated calcium chanel; gamma-2 subunit;
KW   gamma-3 subunit; CACNGLIKE2; neurological disorder; epilepsy; stroke;
KW   head trauma; migraine; affective disorder; depression; anxiety;
KW   schizophrenia; neurodegenerative disorder; Alzheimer's disease;
KW   cognitive disorder; pain; cancer; vaccine; expressed sequence tag; EST;
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..945
FT                   /*tag= a
XX
PN   WO200014225-A1.
XX
PD   16-MAR-2000.
XX
PF   06-SEP-1999;   99WO-GB002945.
XX
PR   08-SEP-1998;   98GB-0019589.
XX
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Duckworth DM, Hayes PD;
XX
DR   WPI; 2000-256977/22.
DR   P-PSDB; AAY84375.
XX
PT   CACNGLIKE polynucleotides and polypeptides, useful in diagnostic assays
PT   and for treating conditions such as Alzheimer's.
XX
PS   Claim 11; Page 31; 33pp; English.
XX
CC   The present sequence represents an expressed sequence tag (EST)
CC   corresponding to a putative human neuronal voltage-gated calcium chanel
CC   gamma-2 and gamma-3 subunits, designated CACNGLIKE2 polypeptides. The
CC   CACNGLIKE2 polynucleotide and polypeptide are useful for treating
CC   neurological disorders, epilepsies, stroke, head trauma, migraine,
CC   affective disorders including depression and anxiety, schizophrenia,
CC   neurodegenerative disorders including Alzheimer's disease, cognitive
CC   disorders, pain and cancer. They may also be used to configure screening
CC   methods for detecting the effect of added compounds on the production of
CC   mRNA in cells. The polypeptides may also be used as vaccines to induce an
CC   immunological response in mammals
XX
SQ   Sequence 945 BP; 232 A; 274 C; 217 G; 222 T; 0 U; 0 Other;

RESULT 13
AA29793
ID   AA29793 standard; DNA; 948 BP.
XX

Query Match     7.9%;  Score 52;  DB 3;  Length 945;
Best Local Similarity   65.5%;             Pred. No. 0.0015;
Matches  76;  Conservative   0;  Mismatches  40;  Indels   0;  Gaps   0;

Qy   218 GTGTGAAAAGGGCGTTCAGGTACTACTGACCACCACATCGGCGGCCTTCGCGGCTTTGGCCT 277
         |||||   || |||||||||||||||||||| ||||||||||||| || ||||||||||||||
Db     9 GTGTGACAGAGGTATCCAGATGTTGATCACCACTGGTAGGAGCCTTGCCGGCTTTTAGTTT 68

Qy   278 CATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGCTCTCATCTGCA 333
         ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db    69 AATGACCATTGCAGTGGGCACGGACTACTGGTGTATATTCCAGAGGTGTGTGCAGGA 124
```

```
AC   AA299793;
XX
DT   12-JUL-2000  (first entry)
XX
DE   DNA encoding a neuronal voltage-gated calcium chanel.
XX
KW   Human; neuronal voltage-gated calcium chanel; gamma-2 subunit;
KW   gamma-3 subunit; CACNGLIKE2; neurological disorder; epilepsy; stroke;
KW   head trauma; migraine; affective disorder; depression; anxiety;
KW   schizophrenia; neurodegenerative disorder; Alzheimer's disease;
KW   cognitive disorder; pain; cancer; vaccine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..948
FT                   /*tag= a
FT                   /product= "voltage-gated calcium chanel"
XX
PN   WO200014225-A1.
XX
PD   16-MAR-2000.
XX
PF   06-SEP-1999;   99WO-GB002945.
XX
PR   08-SEP-1998;   98GB-0019589.
XX
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Duckworth DM, Hayes PD;
XX
DR   WPI; 2000-256977/22.
DR   P-PSDB; AAY84374.
XX
PT   CACNGLIKE polynucleotides and polypeptides, useful in diagnostic assays
PT   and for treating conditions such as Alzheimer's.
XX
PS   Claim 4; Page 29; 33pp; English.
XX
CC   The present sequence encodes a putative human neuronal voltage-gated
CC   calcium chanel gamma-2 and gamma-3 subunits, designated CACNGLIKE2
CC   polypeptides. The CACNGLIKE2 polynucleotide and polypeptide are useful
CC   for treating neurological disorders, epilepsies, stroke, head trauma,
CC   migraine, affective disorders including depression and anxiety,
CC   schizophrenia, neurodegenerative disorders including Alzheimer's disease,
CC   cognitive disorders, pain and cancer. They may also be used to configure
CC   screening methods for detecting the effect of added compounds on the
CC   production of mRNA in cells. The polypeptides may also be used as
CC   vaccines to induce an immunological response in mammals
XX
SQ   Sequence 948 BP; 234 A; 273 C; 218 G; 223 T; 0 U; 0 Other;

RESULT 14
ABA09004
ID   ABA09004 standard; cDNA; 1383 BP.
XX
AC   ABA09004;
XX
DT   11-JAN-2002  (first entry)
XX

Query Match     7.9%;  Score 52;  DB 3;  Length 948;
Best Local Similarity   65.5%;             Pred. No. 0.0015;
Matches  76;  Conservative   0;  Mismatches  40;  Indels   0;  Gaps   0;

Qy   218 GTGTGAAAAGGGCGTTCAGGTACTACTGACCACCACATCGGCGGCCTTCGCGGCTTTGGCCT 277
         |||||   || |||||||||||||||||||| ||||||||||||| || ||||||||||||||
Db     9 GTGTGACAGAGGTATCCAGATGTTGATCACCACTGGTAGGAGCCTTGCCGGCTTTTAGTTT 68

Qy   278 CATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGCTCTCATCTGCA 333
         ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db    69 AATGACCATTGCAGTGGGCACGGACTACTGGTGTATATTCCAGAGGTGTGTGCAGGA 124
```

DE   Human voltage gated C1 channel subunit homologue cDNA, SEQ ID NO:780.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer; 88.
XX
OS   Homo sapiens.
XX
PN   WO20015718B-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US003800.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT; Liu C, Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   P-PSDB; ABB11760.
XX
PT   Human proteins and DNA encoding sequences useful for preventing, treating
PT   or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT   and cancer.
XX
PS   Claim 1; Page 691-692; 1963pp; English.
XX

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides,
antibodies against the polypeptides, methods of detecting the nucleotides
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; tissue growth activity;
haematopoiesis regulatory activities; stem cell growth factor activity;
immunomodulatory activity; activity; activin- or inhibin-related activities;
chemotactic or chemokinetic activities; haemostatic, thrombotic or
thrombolytic activities; receptor or ligand activities; or may be
involved in oncogenesis, cancer cell proliferation or metastasis.
Depending on their biological activities, polypeptides and nucleotides of
the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used in
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness.

CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a cDNA encoding a
CC   novel human polypeptide of the invention
XX
SQ   Sequence 1383 BP; 336 A; 407 C; 341 G; 299 T; 0 U; 0 Other;

Query Match   7.9%; Score 52; DB 4; Length 1383;
Best Local Similarity   65.5%;
Matches   76; Conservative   0; Mismatches   40; Indels   0; Gaps   0;

Qy   218 GTGTGAAAAGGCGGTTCAGGTACTACTGACCACCATCGGGGCCCTTCGCGGCGGCTTTTGGCCT 277
          ||||| | || ||| | | |||| | ||||| || || || | ||||| || ||||
Db   429 GTGTGACAGAGGTATCCAGATGTTGATCACCACTGTAGGAGCCTTGCCGCTTTTAGTTT 488

Qy   278 CATCGACCATCGCCATCACAGCACTGACTACTGGGCTTCTACACAAAGAGCTCTCATCTGCA 333
          ||| |||| ||| | || || | ||||| ||| |||| | | | ||||| | || |
Db   489 AATGACCAATGCCAGTGGGGCACGGACTTATTATTCCAGAGAGGTGTGTGCCAGGA 544

RESULT 15
AAC77216
ID   AAC77216 standard; cDNA; 1854 BP.
XX
AC   AAC77216;
XX
DT   08-FEB-2001 (first entry)
XX
DE   Human ORFX ORF2771 polynucleotide sequence SEQ ID NO:5541.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO20005473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US008621.
XX
PR   31-MAR-1999; 99US-0127607P.
PR   02-APR-1999; 99US-0127636P.
PR   05-APR-1999; 99US-0127728P.
PR   30-MAR-2000; 2000US-00540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA, Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB43007.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease.
XX
PS   Claim 5; Page 4724-4725; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX

CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC   anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC   cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC   dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC   antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC   sequences can be used for determining the presence of or predisposition
CC   to, or preventing or treating pathological conditions associated with an
CC   ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC   proteins in gene therapy vectors. The proteins and nucleic acids may be
CC   used to treat cancers, proliferative disorders, neurodegenerative
CC   disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC   diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC   storage, systemic lupus erythematosus, severe combined immunodeficiency
CC   (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC   disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC   cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC   enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ   Sequence 1854 BP; 460 A; 549 C; 414 G; 429 T; 0 U; 2 Other;

Query Match   7.9%; Score 52; DB 3; Length 1854;
Best Local Similarity   65.5%;
Matches   76; Conservative   0; Mismatches   40; Indels   0; Gaps   0;
Pred. No. 0.0019;

Qy   218 GTGTGAAAAGGGGCGGTTCAGGTACTACTGACCACCATCCGGCGCCTTCCGCGGCTTTTGGCCT 277
             ||||||||| ||| |||||||| || || ||||| |||||||||| ||| |||||||| |||||
Db   354 GTGTGACAGAGGTTATCCAGATGTTGATCACCACCTGTAGGAGACCCTTTGCCGCTTTTAGTTT 413

Qy   278 CATGACCATCGCCATCAGCACTGACTACTGGCTCTCACACAAGAGCTCTCACTGCA 333
             |||||| || |||||||| ||| |||| ||||| ||| |||||| |||||||||||
Db   414 AATGACCATTGCAGTGGGCACCGACTACTGGGTTTATATTCCAGAGAGGTGTGTGCAGGA 469

Search completed: May 22, 2004, 15:44:52
Job time : 307 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 22, 2004, 15:32:57 ; Search time 2367 Seconds
                 (without alignments)
                 8276.126 Million cell updates/sec

Title: US-10-060-066-2
Perfect score: 656
Sequence: 1 tatggcttcactctgcacct..........tgtatttcttattcttcttg 656

Scoring table: IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est3:*
                 13: gb_est4:*
                 14: gb_est5:*
                 15: em_estfun:*
                 16: em_estom:*
                 17: em_estfun:*
                 18: em_gss_hum:*
                 19: em_gss_inv:*
                 20: em_gss_pln:*
                 21: em_gss_vrt:*
                 22: em_gss_fun:*
                 23: em_gss_mam:*
                 24: em_gss_mus:*
                 25: em_gss_pro:*
                 26: em_gss_rod:*
                 27: em_gss_phg:*
                 28: em_gss_vrl:*
                 29: gb_gss1:*
                     gb_gss2:*

	No.	Score				ID	Description
c	5	234.2	35.7	558	13	BQ266161	BQ266161 NISC_ff11
c	6	189.6	28.9	547	10	BE102334	BE102334 UI-R-BT1-
c	7	174	26.5	245	10	BE864111	BE864111 UI-M-BH1-
c	8	159.8	24.4	663	10	BB641732	BB641732 BB641732
c	9	155.8	23.8	472	9	AW060348	AW060348 UI-M-BH3-
	10	121.2	18.5	169	29	CG620013	CG620013 OST316577
	11	102.2	15.6	983	29	CN50489W	CN50489W Tetraodon
	12	92	14.0	655	14	CD215523	CD215523 pgp2n_pk0
	13	86.4	13.2	528	14	CK332731	CK332731 H82128L2-
	14	86.4	13.2	600	12	BG803929	BG803929 0243-54_M
	15	86.4	13.2	961	10	CF583776	CF583776 AGENCOURT
	16	83.6	12.7	600	13	BU925588	BU925588 7124-20_M
	17	81.8	12.5	594	12	BI981765	BI981765 fu50dll.y
	18	81.4	12.4	587	12	BI844628	BI844628 fg14g05.y
c	19	74.4	11.3	999	29	CN504NDM	CN504NDM Tetraodon
	20	73.6	11.2	691	12	BG912818	BG912818 60280T258
	21	73.6	11.2	697	13	BY733596	BY733596 BY733596
	22	73.2	11.2	592	29	FR0019468	FR0019468 F.rubripe
c	23	62	9.5	615	29	CNS04174	CNS04174 Tetraodon
	24	61.4	9.4	687	29	CNS04LQO	CNS04LQO Tetraodon
c	25	61.4	9.4	688	10	BE212599	BE212599 IpBrn0032
c	26	61.4	9.4	697	10	BE574164	BE574164 IpBrn0116
	27	61.4	9.4	877	14	CA496554	CA496554 AGENCOURT
	28	58.2	8.9	771	14	CA750036	CA750036 UI-M-FD0-
c	29	56	8.5	923	29	CNS021DK	CNS021DK Tetraodon
c	30	54.4	8.3	599	9	FR0019451	FR0019451 F.rubripe
c	31	54.4	8.3	580	10	BE222518	BE222518 hv92c08.x
c	32	54.4	8.3	599	9	AI915567	AI915567 wellb12.x
	33	54.4	8.3	606	10	AI680925	AI680925 tx42h02.x
c	34	54.4	8.3	640	10	BE670632	BE670632 7e37e05.x
	35	54	8.3	2427	11	AK045362	AK045362 Mus muscu
	36	54	8.2	1043	29	CNS03MBE	CNS03MBE Tetraodon
c	37	53.6	8.2	514	13	BX338963	BX338963 BX338963
	38	53.6	8.2	649	12	BM951619	BM951619 UI-M-EG0-
	39	53.6	8.2	855	12	BI739481	BI739481 603361722
c	40	53.6	8.2	948	29	AY411029	AY411029 Mus muscu
c	41	53.6	8.2	1103	13	BX403654	BX403654 BX403654
c	42	53.4	8.1	157	9	AI693801	AI693801 wd87a03.x
	43	53.4	8.1	965	12	BI457751	BI457751 60319B193
c	44	53	8.1	1345	28	CC273194	CC273194 CH261-37J
	45	52.8	8.0	477	10	BE861331	BE861331 UI-M-AK0-

ALIGNMENTS

RESULT 1
CG591004
LOCUS   CG591004   656 bp   DNA   linear   GSS 02-OCT-2003
DEFINITION OST245090 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST245090, genomic survey sequence.
ACCESSION   CG591004
VERSION   CG591004.1 GI:37398453
KEYWORDS   GSS.
SOURCE   Mus musculus
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 656)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlnauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE   Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT   Contact: Zambrowicz BP
            Lexicon Genetics Incorporated
            OmniBank
            4000 Research Forest Drive, The Woodlands, TX 77381, USA

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	644	98.2	656	29	CG591004	CG591004 OST245090
2	474	72.3	504	29	CG601451	CG601451 OST273800
3	452	68.9	478	29	CG665024	CG665024 OST452876
4	332.4	50.7	337	10	BE647856	BE647856 UI-M-BH1-

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.

```
FEATURES             Location/Qualifiers
     source          1..656
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/Ev"
                     /db_xref="taxon:10090"
                     /clone="OST245090"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129SV/Ev"
ORIGIN

Query Match         98.2%; Score 644; DB 29; Length 656;
Best Local Similarity   100.0%; Pred. No. 3.3e-122;
Matches   656;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TATGGCTTCACTCTGCACCTTGCCTTCTCGCCTGCCACCTGTGCCCCGCCCCCCTC   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TATGGCTTCACTCTGCACCTTGCCTTCTCGCCTGCCACCTGTGCCCCGCCCCCCTC   60

Qy   61 TGGCCCCTTGAGGCGCCCCCCCCCAGCTTCTCGCCTGTGATCCCCCAGCCGCCACGG  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGGCCCCTTGAGGCGCCCCCCCCCAGCTTCTCGCCTGTGATCCCCCAGCCGCCACGG  120

Qy  121 CCCCGGCCTCCGGTCGCCCGGTGCGTGGCCCACCGGCCGCCCCCCCGGGCCGGCACGG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CCCCGGCCTCCGGTCGCCCGGTGCGTGGCCCACCGGCCGCCCCCCCGGGCCGGCACGG 180

Qy  181 AGTCATTGAAACGCTGAAATGAAGAGAGAGGGTTTGTGGTGTGAAAAGGGCGTTCAGGTAC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AGTCATTGAAACGCTGAAATGAAGAGAGAGGGTTTGTGGTGTGAAAAGGGCGTTCAGGTAC 240

Qy  241 TACTGACCACCATCGGGCGCCTTGCGGCCTTTGGCGGTTCATGACCATCGCCATCAGCACTG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TACTGACCACCATCGGGCGCCTTGCGGCCTTTGGCGGTTCATGACCATCGCCATCAGCACTG 300

Qy  301 ACTACTGGCTCTACACAAGAGCTCTCATCTGCACAACACCACCTCACAGCAGGTGATG 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ACTACTGGCTCTACACAAGAGCTCTCATCTGCACAACACCACCTCACAGCAGGTGATG 360

Qy  361 ACGGACCACCCCCATCGTGGGGGGCCAGTGGCCTCCTCCGAGAAGAAGAAGGACCCTGGGGCACGG 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ACGGACCACCCCCATCGTGGGGGGCCAGTGGCCTCCTCCGAGAAGAAGAAGGACCCTGGGGCACGG 420

Qy  421 CACATTCAGGCCTTCGGCGGATATGCTGGCGCCTGGAAGGTAGGGTGCAGGAAGGCCCTGGGT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CACATTCAGGCCTTCGGCGGATATGCTGGCGCCTGGAAGGTAGGGTGCAGGAAGGCCCTGGGT 480

Qy  481 GGCCACCTCCGNTGAACGCCTGCCTNCAGCAAATCGACTTNTAGTGGTGCAGGAAGGCCCTGGGT 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GGCCACCTCCGNTGAACGCCTGCCTNCAGCAAATCGACTTNTAGTGGTGCAGGAAGGCCCTGGGT 540

Qy  541 CCCAAGGNCTTGAGGGGTGTGCTTTACTGCTCGGAGGCGGGNTNATTCCATGCACCCCTCCA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CCCAAGGNCTTGAGGGGTGTGCTTTACTGCTCGGAGGCGGGNTNATTCCATGCACCCCTCCA 600

Qy  601 TNCCTNTNTACAAATCCCTCTTCTGCGGGGATTNTNGNTCGTTATTTNTTTTATTTTCTTTG 656
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TNCCTNTNTACAAATCCCTCTTCTGCGGGGATTNTNGNTCGTTATTTNTTTTATTTTCTTTG 656
```

RESULT 2

```
LOCUS       CG601451              504 bp    DNA     linear   GSS 02-OCT-2003
DEFINITION  Mus musculus 129SV/Ev Mus musculus genomic clone
            OST273800, genomic survey sequence.
ACCESSION   CG601451
VERSION     CG601451.1  GI:37420111
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 504)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q, Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     OmniBank
            Contact: Zambrowicz BP
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1..504
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/Ev"
                     /db_xref="taxon:10090"
                     /clone="OST273800"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129SV/Ev"
ORIGIN

Query Match         72.3%; Score 474; DB 29; Length 504;
Best Local Similarity   99.8%; Pred. No. 3e-87;
Matches   485;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   1;

Qy    1 TATGGCTTCACTCTGCACCTTG-ACCTTGCCCTCTTGCAGCCGCCACCTGTCGCCCCGCCCCCCTC   60
        |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   15 TATGGCTTCACTCTGCACCTTGCACCTTGCCCTCTTGCAGCCGCCACCTGTCGCCCCGCCCCCCTC   73

Qy   61 TGGCCCCTTGAGGCGCCCCCCCCCAGCTTCTCGCCTGTGATCCCCCAGCCGCCACGG  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74 TGGCCCCTTGAGGCGCCCCCCCCCAGCTTCTCGCCTGTGATCCCCCAGCCGCCACGG  133

Qy  121 CCCCGGCCTCCGGTGCCCCGGTGGTGGCCCCACCGGCCGCCCCCCGGGCCGGCACGG 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  134 CCCCGGCCTCCGGTGCCCCGGTGGTGGCCCCACCGGCCGCCCCCCGGGCCGGCACGG 193

Qy  181 AGTCATTGAAACGCTGAAATGAAGAGAGAGAGGGTTTGTGGTGTGAAAAGGGCGTTCAGGTAC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  194 AGTCATTGAAACGCTGAAATGAAGAGAGAGAGGGTTTGTGGTGTGAAAAGGGCGTTCAGGTAC 253

Qy  241 TACTGACCACCATCGGGCGCCTTCGGCGGCTTTTGGCCCTCATGACCATCGCCATCAGCACTG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  254 TACTGACCACCATCGGGCGCCTTCGGCGGCTTTTGGCCCTCATGACCATCGCCATCAGCACTG 313

Qy  301 ACTACTGGCTCTACACAAGAGCTCTCATCTGCAACACACCACCTCACAGCAGGTGATG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  314 ACTACTGGCTCTACACAAGAGCTCTCATCTGCAACACACCACCTCACAGCAGGTGATG 373

Qy  361 ACGGACCACCCCCATCGTGGGGCCAGTGGCCTCCTCCGAGAAGAAGAAGGACCCCTGGGGCACGG 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  374 ACGGACCACCCCCATCGTGGGGCCAGTGGCCTCCTCCGAGAAGAAGAAGGACCCCTGGGGCACGG 433

Qy  421 CACATTCAGGCCTTCGGCGGATATGCTGGCGCCTGGAAGGTAGGGTGCAGGAAGGCCCTGGGT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  434 CACATTCAGGCCTTCGGCGGATATGCTGGCGCCTGGAAGGTAGGGTGCAGGAAGGCCCTGGGT 493

Qy  481 GGCCAC  486
        ||||||
Db  494 GGCCAC  499
```

```
RESULT 3
CG665024
LOCUS       CG665024                478 bp    DNA    linear   GSS 02-OCT-2003
DEFINITION  OST452876 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST452876, genomic survey sequence.
ACCESSION   CG665024
VERSION     CG665024.1  GI:37488873
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 478)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1..478
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/EV"
                     /db_xref="taxon:10090"
                     /clone="OST452876"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

Query Match             68.9%;  Score 452;  DB 29;  Length 478;
  Best Local Similarity   99.8%;
  Matches  463;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy         1 TATGGCTTCACTCTGCACCTTCGCCTCTCCAGCCACCTGCTCCCCGGCGCCCCCTC 60
             ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        16 TATGGCTTCACTCTG-ACCTTCGCCTCTCCAGCCACCTGCTCCCCGGCGCCCCCTC 74

Qy        61 TGGCCCCTTGAGGCGCCCCCCAGCTTCTCGTGTGTGTGATCCCCCGCCGGCCACGG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        75 TGGCCCCTTGAGGCGCCCCCCAGCTTCTCGTGTGTGTGATCCCCCGCCGGCCACGG 134

Qy       121 CCCCGGCTTCCGCTCCGTCCTCCCCGGTGGTGGCCCACGGCGCCCGGTTGAAACTGG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       135 CCCCGGCTTCCGCTCCGTCCTCCCCGGTGGTGGCCCACGGCGCCCGGTTGAAACTGG 194

Qy       181 AGTCATTGAAACCTGGAATGAAGAGAGGGGGTTTGTGGTGTGAAAAGGCGGTTCAGGTAC 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       195 AGTCATTGAAACCTGGAATGAAGAGAGGGGGTTTGTGGTGTGAAAAGGCGGTTCAGGTAC 254

Qy       241 TACTGACCACCATCGGCGCCTTCGGCGCCTTTTGGCCTCATGACCATCGCCATCAGCACTG 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       255 TACTGACCACCATCGGCGCCTTCGGCGCTTTTGGCCTCATGACCATCGCCATCAGCACTG 314

Qy       301 ACTACTGGCTTCTACACAAGAGCTTCATCTGCAACACAACCTCACAGCAGGTGATG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       315 ACTACTGGCTTCTACACAAGAGCTTCATCTGCAACACAACCTCACAGCAGGTGATG 374

Qy       361 ACGGACACCACCCATCGTGTGGGGGGGGCAGTGGCTTCCTCCGAGAAGAGAGAGACCCTGGGGGGGGCCTCA 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       375 ACGGACACCACCCATCGTGTGGGGGGGGCAGTGGCTTCCTCCGAGAAGAGAGAGACCCTGGGGGGGGCCTCA 434
```

```
Qy       421 CACATTCAGGCGCCTCTGGCGGATATGCTGCCTGGAAGGTAGGGTG 464
             ||||||||||||||||||||||||||||||||||||||||||||||
Db       435 CACATTCAGGCGCCTCTGGCGGATATGCTGCCTGGAAGGTAGGGTG 478

RESULT 4
BB647856
LOCUS       BB647856                337 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  UI-M-BH1-anh-c-12-0-UI.r1 NIH_BMAP_M_S2 Mus musculus cDNA clone
            UI-M-BH1-anh-c-12-0-UI 5', mRNA sequence.
ACCESSION   BB647856
VERSION     BB647856.1  GI:9973676
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 337)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mESTs@mail.nih.gov
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
            should be noted that Bento Soares is generating a small number of
            additional specialized non-redundant arrays of BMAP cDNAs whose
            availability will be considered under appropriate and limited
            collaborative arrangements The following repetitive elements were
            found in this cDNA sequence: 88-148, >GC_rich#Low_complexity
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..337
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH1-anh-c-12-0-UI"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NIH_BMAP_M_S2"
                     /note="Vector: pT7T3D-Pac (pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_M_S1 library, which in turn is a subtracted library
                     derived from a mixture of normalized libraries derived from ten
                     regions of the mouse brain (cerebellum, brain stems,
                     olfactory bulbs, hypothalamus, cortex, amygdala, basal
                     ganglia, pineal gland, striatum, hipoccampus). The driver
                     used for subtraction consisted of a pool of 5,000 clones
                     from the NIH_BMAP_M_S1 library and a pool of 2,000 clones
                     obtained from non-normalized and normalized mouse brain
                     spinal cord libraries."
ORIGIN

Query Match             50.7%;  Score 332.4;  DB 10;  Length 337;
  Best Local Similarity   99.7%;
  Matches  333;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        17 ACCTTCGCCTCTCCAGCCACCTGCTCCCCGGCGCCCCCTCTGGCCCCTTGAGGCCC 76
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4 ACCTTCGCCTCTCCAGCCACCTGCTCCCCGGCGCCCCCTCTGGCCCCTTGAGGCCC 63
```

RESULT 5

```
LOCUS       BQ266161/c
DEFINITION  BQ266161
            NISC_ffilc05.x1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:2865416
            3', mRNA sequence.
ACCESSION   BQ266161                558 bp    mRNA    linear    EST 07-MAY-2002
VERSION     BQ266161.1  GI:20491226
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 558)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen/Robin Humphreys
            cDNA Library Preparation: Life Technologies
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            MGI:1046188
            Plate: LLAM7069  row: F  column: 9
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1..558
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2865416"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /clone_lib="NCI_CGAP_Mam5"
                     /lab_host="DH10B"
ORIGIN

Query Match             35.7%;  Score 234.2;  DB 13;  Length 558;
Best Local Similarity   89.2%;  Pred. No. 6.5e-38;
Matches  256;  Conservative   0;  Mismatches 30;  Indels  1;  Gaps  1;

Qy         77 CCCAGCTTCTGCCTGTGTCGTGTGATCCCCCAGCGTCACGGCGCACGGGCACCTTCCGCTGCC 136
              |||||| ||||||| ||||| |||||| ||||| ||||| |||| ||| |||||||||||||
Db         64 CCCAGCTTCTGCCTGTGTCGTGTCATCCCCCAGCGTCACGGCGCACGGGCTTCCGCTGCC 123
Qy        137 CCGGTGGTTGGGCCCCACGCCCCCGGTTGCCAGTTGGTGAAGTCATTGAAACCGCTG 196
              ||||| |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        124 CCGGTGGTTGGGCCCCACGGGCCCCGGTTGCCAGTTGGTGAAAACTGGAGTCATTGAAACCGCTG 183
Qy        197 GAAATGAAGAGGAGGGGGTTTGTGTGTGAAAAGGCGGTTCAGGTTACTACTGACCACCATCGG 256
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GAATGAAGAGGAGGGGGTTTGTGTGGTGTGAAAAGGGGGTTCAGGTTACTACTGACCACCATCGG 243
Qy        257 CGCCTTCGCGGTGTTTGGGCCTATGACCAATCGGCCATAGCACTGACCTGGCTCTACAC 316
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        244 CGCCTTCGCGGTGTTTGGGCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACAC 303
Qy        317 AAGAGCTCTCATCTGCAACACCACCACCTCACA 350
              ||||||||||||||||||||||||||||||||||
Db        304 AAGAGCTCTCATCTGCAACACCACCACCTCACA 337
```

RESULT 6

```
LOCUS       BE102334/c
DEFINITION  BE102334
            UI-R-BT1-aql-d-02-0-UI.s1 UI-R-BT1 Rattus norvegicus cDNA clone
            UI-R-BT1-aql-d-02-0-UI 3', mRNA sequence.
ACCESSION   BE102334                547 bp    mRNA    linear    EST 13-JUN-2000
VERSION     BE102334.1  GI:8494433
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 547)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to identify it as a clone from the
            normalized corpus-striatum library cDNA Library preparation: M.B.
            Soares Lab Clone distribution: clones will be available through
            Research Genetics (www.resgen.com) The following repetitive
            elements were found in this cDNA sequence: 147-294, >B3#SINE/B2
            PolyA=Yes.
            Seq primer: M13 Forward
FEATURES             Location/Qualifiers
     source          1..547
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-BT1-aql-d-02-0-UI"
                     /dev_stage="adult"
                     /clone_lib="UI-R-BT1"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not I; Site 2: Eco RI; The library
                     UI-R-BT1 is a subtracted library derived from a mixture of
                     the following tissues: hippocampus, thalamus, mid-brain,

Qy        370 CCCATCGTTGGGGCCAGTGCGTCCTTCCCGGAGAGGAGGGACCCTGGGGCACATTCAG 429
              ||||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||
Db        558 CCCATCGTTGGGGCCAGTGCGTCGGCCTTCCCGGAGAAGAGGACCCTGGGGCACATTCAG 499
Qy        430 GCCTTCTGGCCGGATATGCTGTCGGTTGGGAAGTAGGTGCAGGGCACGGCTGCCACTTC 489
              |||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||
Db        498 GCCTTCTGGCCGGATATGTCTGGCCTGGGAAGGTAGGTGCAGGGCACGGCTGCCACTTC 439
Qy        490 CGNTGGAGGCCTGTCGCCTNNCAGCAAAATCGACTNNTAGYGTNGYGTTCGGAGTCTCAAAACCCAAGNC 549
              |||||||||||||||||||||||||||||||| |||||||||||||| |||||||||||||| ||||||
Db        378 CGCTGGAGGCCTGTCGCCTCCAGCAAAATCGACTCTCTAGTGTTCGTGGAGTCTCAAAACCCAAGGTC 379
Qy        550 TTGAGGGGTGGCGTTTACTGGAGGGGGGNTNATTCCATGACACCCCCTCCATNCCTNTNNT 609
              |||||||||||||||||||||| ||||| | ||||||||||||||||||||||||| |||
Db        438 TTGAGGGGTGGCGTTTACTGGAGGGGGTGATTCCATGACACCCCCTCCATCCCTCTGT 319
Qy        610 ACAAATCCCTCTTCTCTGGGGATTNTNGNGNGTTATTTATTTTTCTTTG 656
              ||||||||||||||||||||||||| | | || |||||||||| || |
Db        318 -CCAATCCCTCTTCTCTGGGGATTCTCGTTTATTTTATTTTTTTATACTGG 273
```

```
RESULT 7
LOCUS       BE864111/c
LOCUS       BE864111                 245 bp    mRNA    linear     EST 29-SEP-2000
DEFINITION  UI-M-BH1-anp-a-08-0-UI.r1 NIH_BMAP_M_S2 Mus musculus cDNA clone
            UI-M-BH1-anp-a-08-0-UI 5', mRNA sequence.
ACCESSION   BE864111
VERSION     BE864111.1  GI:10384829
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 245)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      8889548
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD .
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            cDNA Library Preparation: M.B. Soares Lab clone distribution:
            Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
            should be noted that Bento Soares is generating a small number of
            additional specialized non-redundant arrays of BMAP cDNAs whose
            availability will be considered under appropriate and limited
            collaborative arrangements
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..245
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH1-anp-a-08-0-UI"
                     /dev_stage="27-32 days"
```

ORIGIN

```
Query Match     28.9%;  Score 189.6;  DB 10;  Length 547;
Best Local Similarity   90.5%;  Pred. No. 9.7e-29;
Matches  209;  Conservative   0;  Mismatches   21;  Indels    1;  Gaps    1;

Qy        375 CGTGGGGGCGCAGTGCTGGCTCTTCCCGACGAAGAAGGACACCCTGGGGGGCTTCACACATTCAGGCCTC 434
              |||||||||||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||||||
Db        547 CGTGGGGGCGCAGTGCTGGCTCTTCAGAGAGGAAGGACCCCTGGGGGGCTTCACACATTCAGGCCTC 488

Qy        435 TGGCGGGATATGCTGGTCGTCCTCGGAAGGTAGGGGTGGTGGCAGGAAGGCCCCTGGCTGGCCACCTCCGNTG 494
              ||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||| ||
Db        487 TGGCGGGATATGCTGGTCGTCCTCGGAAAGGTAGGGGTGGTGCAGGAAGGCCCCTGGCTGGCCACCTCCGNTG 428

Qy        495 AAGCCTGTGCCCTNCAGCAAATCGACTTNTAGTGTTCTGAGTCTCAACCCAAGGNCTTGAG 554
              |||||||||||||||||||||||||||| ||| |||||||| |||||||||||||||||||
Db        427 AAGCCTGTGCCCTCCAGCAAANCGACTGCTAGTGTGTTCTGAAATCTCAACCCAAAGGTCTTGAG 368

Qy        555 GGGTGTGCTTTACTGCTGGAGGGG-GNTNATTCCATGCACCCCCTCCCATNCC 604
              |||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        367 GGGTGTGCTTTCACTGCTGGAGGAGGACTATGATTCCATGCATCCCTCTGTCC 317
```

```
RESULT 8
LOCUS       BB641732
LOCUS       BB641732                 663 bp    mRNA    linear     EST 26-OCT-2001
DEFINITION  BB641732 RIKEN full-length enriched, 10 days neonate cortex Mus
            musculus cDNA clone A830045A22 5', mRNA sequence.
ACCESSION   BB641732
VERSION     BB641732.1  GI:16476827
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 663)
AUTHORS     Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sasaki,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
            RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL     Unpublished (2001)
TITLE       Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC) Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp;
            URL:http://genome.gsc.riken.go.jp/
COMMENT     Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
            wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
```

ORIGIN

```
Query Match     26.5%;  Score 174;  DB 10;  Length 245;
Best Local Similarity   99.4%;  Pred. No. 1.4e-25;
Matches  174;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps    0;

Qy        283 CCATGCCATCAGCACTGACTACTGGCTCTCACAAGAGAGCTCTCATCTGCAACACCACCA 342
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CCATGCCATCAGCACTGACTACTGGCTCTCACAAGAGAGCTCTCATNTGCAACACCACCA 186

Qy        343 ACCTCACACAGCAGGTGATGACGGACACCACCCCCATCGTGGGGGCCAGTGGGCTCCTCCCGGAGAAGA 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACCTCACACAGCAGGTGATGACGGACACCACCCCCATCGTGGGGGCCAGTGGGCTCCTCCCGAGAAGA 126

Qy        403 AGGACCCTGGGGGGCCTCACACATTCAGGCCCTCTGCGGATATGCTGCCTGGAAGG 457
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 AGGACCCTGGGGGGCCTCACACATTCAGGCCCTCTGCGGGGATATGCTGCCTGGAAGG 71
```

```
    /lab_host="DH10B (life Technologies)"
    /clone_lib="NIH_BMAP_M_S2"
    /note="Vector: pTTT3D-Pac (Pharmacia) with a modified
    polylinker; Site 1: Not I, Site 2: Eco RI; The
    NIH_BMAP_M_S2 library is a subtracted library derived from
    NIH_BMAP_M_S1, which in turn is a subtracted library
    derived from a mixture of normalized libraries from ten
    regions of the mouse brain (cerebellum, brain stems,
    olfactory bulbs, hypothalamus, cortex, amygdala, basal
    ganglia, pineal gland, striatum, hippocampus). The driver
    used for subtraction consisted of a pool of 5,000 clones
    from the NIH BMAP M S1 library and a pool of 2,000 clones
    obtained from non-normalized and normalized mouse brain
    spinal cord libraries."
```

medulla, corpus striatum, cerebral cortex and testis. For
a detailed description of the library from which this
clone was derived, please visit our web site at
rateat.eng.uiowa.edu. The subtraction has been previously
described in (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_TISSUE=corpus-striatum
TAG_LIB=UI-R-BT1
TAG_SEQ=CTAGG"

RESULT 9
AW060348/c

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES   Location/Qualifiers
source   1..663
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="A930045A22"
                     /tissue_type="cortex"
                     /dev_stage="10 days neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 10 days neonate
                     cortex"
                     /note="Site 1: SalI; Site 2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5',
                     GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5', GAGAGAGAGATTCTCGAGTTAAYTAATTAAAATTAATTCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."

ORIGIN

Query Match   24.4%; Score 159.8; DB 10; Length 663;
Best Local Similarity   80.6%;
Matches 199; Conservative   0; Mismatches 47; Indels   1; Gaps 1;

```
Qy      1 TATGGCTTCACTCTGCACCTTCGCCTCTGCAGCCACTCGTCGCTCGCCGGCCCCCTC 60
            |||||||| |||||| |||||||||||||||| ||||||||||||||||||||| |||
Db    409 TATGGCTTTACTCTG-ACCTTTGCCTANGCAGCCACTCGTCGCTCGCCGGCCCACTC 467

Qy     61 TGGCCCCTTGAGGCCCCCCCAGCTTTCTGCCTGTGCTGTGATCCCCCAGCGGCCGCACGG 120
            ||||| ||||||||||||| ||||| || || ||||||||||||||||| |||| |||||
Db    468 TGGCCTCTTGAGGCCCCCCCAGCTTTCGTGTCTGTGATACCCCCAGCCGTTCGCACGG 527

Qy    121 CCCCGCCTCCGCGCCCCGGTGGTGCCCCCACCGGCCCCCCCGGTTGCCAGTGGGTGAAACTGG 180
            |||||||||| |||||||||||| |||||| || ||||| |||||| |||||||||||||||
Db    528 CCCCGCCTCTCTCGGCCGGTGGTGGCCCCACGGCGCCCCATTGGCCGAAACTGG 587

Qy    181 AGTCATTGAAACGCTGGAATGAAGAGAGGAGGGGTTTGTGGTGGTGTGAAAAGGGGGTTCAGGTAC 240
            ||||||   |  ||   |||||||| || ||||| || || ||||| | ||| || |||| |||
Db    588 AGTCATCATCGTACCCCGTAATGAAGGATAGGAGGAGATGTGCTGTGTTACAGGCTTTTATCGCT 647

Qy    241 TACTGAC 247
            || |||
Db    648 TCCCGAC 654
```

LOCUS   AW060348   472 bp   mRNA   linear   EST 04-OCT-1999
DEFINITION UI-M-BH1-anh-c-12-0-UI.s1 NIH_BMAP_M_S2 Mus musculus cDNA clone
            UI-M-BH1-anh-c-12-0-UI 3', mRNA sequence.
ACCESSION   AW060348
VERSION   AW060348.1 GI:6008099
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 472)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE   Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED   8889548
COMMENT   Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mBST@mail.nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to identify it as a clone from the
            normalized hypothalamus library cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
            available by the means that is soon to be determined. When NIH
            determines the means for distribution of the BMAP cDNA clones, this
            record will be updated accordingly when that means is determined.
            The following repetitive elements were found in this cDNA sequence:
            150-287, >B3#SINE/B2
            Seq primer: M13 Forward
            PolyA=Yes.

FEATURES   Location/Qualifiers
source   1..472
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57Bl/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH1-anh-c-12-0-UI"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NIH_BMAP_M_S2"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not I; Site 2: Eco RI; The
                     NIH_BMAP_M_S2 library is a subtracted library derived from
                     NIH_BMAP_M_S1, which in turn is a subtracted library
                     derived from a mixture of normalized libraries from ten
                     regions of the mouse brain (cerebellum, brain stems,
                     olfactory bulbs, hypothalamus, cortex, amygdala, basal
                     ganglia, pineal gland, striatum, hipoccampus). The driver
                     used for subtraction consisted of a pool of 5,000 clones
                     from the NIH_BMAP_M_S1 library and a pool of 2,000 clones
                     obtained from non-normalized and normalized mouse brain
                     spinal cord libraries.
                     TAG TISSUE=hypothalamus
                     TAG LIB=NIH_BMAP_M_S2
                     TAG_SEQ=CGGTA"

ORIGIN

Query Match   23.8%; Score 155.8; DB 9; Length 472;
Best Local Similarity   85.5%;
Matches 177; Conservative   0; Mismatches 29; Indels   1; Gaps 1;

```
Qy    450 CTGGAAGGTAGGGTGCAGGAGGCCCTGCCACTTCCGNTGAAGCCTGTGCCTNCA 509
            |||||||||||||||||||| ||||||||    |||| ||||||||||||||| ||
Db    472 CTGGAAGGTAGGGTGCAGGAAGGCCCTGGCTGGCCACCTTCCGCTGAAGCCTGTGCCTCCA 413
```

RESULT 10

```
LOCUS       CG620013                 169 bp    DNA     linear   GSS 02-OCT-2003
DEFINITION  CG620013 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST316577, genomic survey sequence.
ACCESSION   CG620013
VERSION     CG620013.1  GI:37443862
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 169)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1..169
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/Ev"
                     /clone="OST316577"
                     /db_xref="taxon:10090"
                     /clone_lib="Mus musculus 129Sv/Ev"
```

```
Query Match             18.5%;  Score 121.2;  DB 29;  Length 169;
Best Local Similarity   94.2%;
Matches  145;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

ORIGIN
```

```
Qy    510 GCAAATCGACTTNNTAGTGTTGTTCTCGAGTCTCAACCCAAGGNCTTGAGGGGTGTGCTTTACTG 569
          |||||||||||| |||||||||||||||||||||||| |||||||| |||||||||||| |||||||
Db    412 GCAAATCGACTTCACTTCTAGTGTTCTCGAGTCTCATCCCAAGGTCTTGAGGGGTGCTTTACTG 353

Qy    570 CTGGAGGCCGGNTNATTCCATGCACCCCTCCATNCCTTNTNTNTACNAAATCCCTTCTTCTTCTTGGGGA 629
          ||||||||||| | ||||||||||||||||||| |||| | | ||| |||||||||||||||||||| ||
Db    352 CTGGAGGCCGGTGTGATTCCATGCACCCCTCCATNCCTCTGT-CCAAATCCCTTCTTCTTCTTGGGGAT 294

Qy    630 TTNTNGNGNTGTTATTTNTTTATTTTCTTTG 656
          | | | | || |||| ||||||||||| ||
Db    293 TCTCGTTTTATTTTTATTTTTATACTGG 267
```

RESULT 11

```
LOCUS       CNS0489W                 983 bp    DNA     linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            090G15 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL279005
VERSION     AL279005.1  GI:8016339
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
   PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359937
   PUBMED   10899143
REFERENCE   3  (bases 1 to 983)
  AUTHORS   Roest Crollius,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT     - web : www.genoscope.cns.fr)
            This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..983
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="090G15"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : COBG090AD08LP1-end : T7"
```

```
Query Match             15.6%;  Score 102.2;  DB 29;  Length 983;
Best Local Similarity   67.5%;
Matches  168;  Conservative    0;  Mismatches   63;  Indels   18;  Gaps    2;

ORIGIN
```

```
Qy    214 TGTGGTGTGAAAAGGCGTTCAGGTACTACTGACCACATCGGCGCGCCTTCGCGGCGTTTG 273
          |||||||||||   |||| ||||| || || || |||||||   |||||||||| ||| |
Db    293 TGTGGTGTGAGCGGGGCATCCAGGTGCTGCTGCTCACCACCATGGGAGCGGTTGCAGCCTTCG 352

Qy    274 GCCTCATGACCATGCCATCAGCACTGACTACTGGCTCTACACAAGAGCTCTCATCTGCA 333
          ||| |||||| || |||| ||||| |||||||||||| |||| |||| || || | |||
Db    353 CCCTGATGACGGTGGCCATGGACACCGACTACTGGCTGGCTCGGCCTGGATTCATTGCA 412

Qy    334 ACACCACCTAACCTACAGCAGGTGATGACGGACACACCCATCGTGGGGGGCGGTGGCTCCT 393
          |||| ||   |||| ||||| || || ||   |||||||||||          ||||||||
Db    413 ACACGACGCCCAACTCACCACCAGGACGAC------------------TCCAACAGCAACA 454

Qy    394 CCGACAAGAAGGACCCTGGGGGGCCTTCACACATTCAGGCCTGGGGATATGCTGCCTGG 453
          || ||||||||||  || ||||||| ||  |||| ||||| ||  ||| |||| |||||
Db    455 AAGACAAGAAGGAGGACCCCGGGGGCCCCTCCCCACTCGGGCCGGAGGAGGAATCTGCTCCTGG 514

Qy    454 AAGGTAGGG 462
```

RESULT 12

```
LOCUS       CD215523                 655 bp    mRNA    linear    EST 20-MAY-2003
DEFINITION  pgp2n.pk008.b11 Normalized chicken pituitary/hypothalamus/pineal
            cDNA library (pgp2n) Gallus gallus cDNA clone pgp2n.pk008.b11 5'
            similar to gb|AAK72510.1 (AY037891) calcium channel gamma 4 subunit
            [Gallus gallus], mRNA sequence.
ACCESSION   CD215523
VERSION     CD215523.1  GI:30955522
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Porter,T.E. and Cogburn,L.A.
  TITLE     Chicken ESTs from pituitary/hypothalamus/pineal gland
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1..655
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Commercial broiler chickens"
                     /clone="pgp2n.pk008.b11"
                     /sex="Male and Female"
                     /tissue_type="Pituitary Gland/Hypothalamus/Pineal Gland"
                     /dev_stage="Embryonic (d12,d14,d19); post-hatch (1,3,5,7,9
                     weeks)"
                     /lab_host="E. Coli EMDH10B"
                     /clone_lib="Normalized chicken
                     pituitary/hypothalamus/pineal cDNA library (pgp2n)"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each tissue at different
                     ages. Single pass sequencing from 5'-end"
ORIGIN
```

```
Db          515 AAGGTACGG 523
            |||||| |||
Qy          515 AAGGTACGG 523
```

Query Match   14.0%; Score 92; DB 14; Length 655;
Best Local Similarity   65.9%;
Matches 162; Conservative   0; Mismatches 60; Indels 24; Gaps 1;

```
Qy    214 TGTGGGTGTGAAAAGGGCGGTTCAGGTACTACTGACCACCATCGCGCCTTCGGGGCTTTTG 273
          ||||| |||| ||  |||||| ||||||||| || |||||   || ||||||| || ||
Db     37 TGTGGTGTGACCGTGGGGTCGCAGAATGCTGCTGCCACGGGTGGAGCTTTTCCGCGCTTCA 96

Qy    274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTTCACAAGAGCTCTCATCTGCA 333
          ||||||||||||||||||||||||||||||||||||||| || |||||| ||||||||||
Db     97 GCCTCATGACCATCGCCATCGGCCATCGGCGTACCCACTACTGGCTCTTCACGGGCACATCGCA 156

Qy    334 ACACCACCAACCTCACAGCAGTGATGACGGACGCACCCATGTGGGGGGCAGTGGCTCCT 393
          ||||||||||||||||||||||||| |||||| ||||||||||||||||||| |||||||
Db    157 ACGGCACCAACATCACGACGGACGACTACGGACGCAGGGGCCCC 195

Qy    394 CCGAGAAGAAGGACGACCCTGGGGGGCCTCACACAATTCAGGCGCGGGGGCGGAATATCCTCGGTGG 453
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||   ||
Db    196 ---CGGAGAAAACGCGAGCGGGGGCGGACCTCACGCACGGGGGGCCTCTTCGGAGGATCTCGCTGCTTCG 252

Qy    454 AAGGTA 459
          |||| |
Db    253 AAGGAA 258
```

RESULT 13

```
LOCUS       CK332731                 528 bp    mRNA    linear    EST 19-DEC-2003
DEFINITION  H8212B12-5 NIA Mouse Unique Gene Set Version 2 Mus musculus cDNA
            clone H8212B12 5', mRNA sequence.
ACCESSION   CK332731
VERSION     CK332731.1  GI:40232323
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G.,
            Martin,P.R., Stagg,C.A., Bassey,U., Aiba,K., Hamatani,T.,
            Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.
  TITLE     Assembly, verification, and initial annotation of NIA 7.4K mouse
            cDNA clone set
  JOURNAL   Genome Res. 12 (12), 1999-2003 (2002)
  MEDLINE   22354164
  PUBMED    12466305
  COMMENT   Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: H8212 row: B column: 12
            Seq primer: M13 Reverse
            High quality sequence stop: 528
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..528
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="niaEST:H8212B12-5"
                     /db_xref="taxon:10090"
                     /clone="H8212B12"
                     /sex="mixed"
                     /dev_stage="mixed"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse Unique Gene Set Version 2"
                     /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
                     NotI; Mouse cDNA project by the Laboratory of Genetics,
                     National Institute on Aging (NIA), Intramural Research
                     Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This
                     clone is among a rearrayed set of 11,424 clones from more
                     than 20 cDNA libraries."
ORIGIN
```

Query Match   13.2%; Score 86.4; DB 14; Length 528;
Best Local Similarity   66.8%;
Matches 123; Conservative   0; Mismatches 61; Indels 0; Gaps 0;

```
Qy    214 TGTGGGTGTGAAAAGGGCGGTTCAGGTACTACTGACCACCATCGCGCCTTCGGGGCTTTTG 273
          ||||| |||| ||  |||||| |||||||||||| || |||||   || ||||||| ||
Db    214 TGCGGATGCGACCGCGGGGGCTGCGCAGATGCTGCTGCCACGGGTGCTGACCACGGGGCGCTTTGT 273

Qy    274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTTCACAAGAGCTCTCATCTGCA 333
          |||||||||||||||||||||||||||||| |||||||||| || |||||| ||||||||
Db     98 GCCTCATGACCATCGCCATCGGCCATCGGCGTACCCACTACTGGCTCTTCACGGGCACATCGCA 157

Qy    334 ACACCACCAACCTCACAGCAGTGATGACGGACGCACCCATGTGGGGGGCAGTGGCTCCT 393
          |||||||||||||||||||||||| |||||| ||||||||||||||||||| |||||||
Db    158 CGCTCATGGCCATCGGCGTACCCACTACTGGCTCTTCACGGGCACGCCACCATCTGCA 217

Qy    394 CCGA 397
          ||||
Db    218 ACGGCACCAACCTGACCATGGACGACGACGGGCCCCGGGCGCTCGGCGGACCTCA 277

Db    278 CCCA 281
```

RESULT 14
BG803929

```
LOCUS       BG803929      600 bp    mRNA    linear   EST 20-DEC-2001
DEFINITION  0243-54 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,
            mRNA sequence.
ACCESSION   BG803929
VERSION     BG803929.1  GI:17950842
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 600)
  AUTHORS   Mu,X., Zhao,S., Pershad,R., Hsieh,T.-F., Scarpa,A., Wang,S.W.,
            White,R.A., Beremand,P.D., Thomas,T.L., Gan,L. and Klein,W.H.
  TITLE     Gene expression in the developing mouse retina by EST sequencing
            and microarray analysis
  JOURNAL   Nucleic Acids Res. 29 (24), 4983-4993 (2001)
  MEDLINE   21671825
   PUBMED   11812828
COMMENT     Contact: Klein WH
            Department of Biochemistry and Molecular Biology
            University of Texas M.D. Anderson Cancer Center
            Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
            Tel: 713 792 3646
            Fax: 713 790 0329.
FEATURES             Location/Qualifiers
     source          1..600
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="neural retina"
                     /dev_stage="embryonic day 14.5 post-fertilization"
                     /clone_lib="Mouse E14.5 retina lambda ZAP II library"
ORIGIN

Query Match         13.2%;  Score 86.4;  DB 12;  Length 600;
Best Local Similarity  66.8%;  Pred. No. 1.7e-07;
Matches  123;  Conservative  0;  Mismatches  61;  Indels  0;  Gaps  0;

Qy    214 TGTGGTGTGAAAAGGGGCGTTCAGGTACTACTGACCACCATCGGCGCCCTTCGCGGCGTTTTG 273
              ||||| ||   ||   |||       || || || || ||    ||    ||     ||
Db     91 TGCCAATGCGACCCGGCGGGCCTGCAGGATGCTCGCTGACCACGGCCCGGAGCCTTCGGCGCCGCCTTCT 150

Qy    274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGACTCTCATCTGCA 333
              |||||| |||||| |||| |  || ||||||||||| ||| |  ||   || ||||||
Db    151 CGCTCATGGCCATCGGCATCGGCGACTACTGGCTGTACTCCAGCGGCCACATCTGCA 210

Qy    334 ACACCACCAACCTCACAGCAGGTGATGACGGACCACACCCATCGTGGGGGGGCCAGTGGCTCCT 393
              || ||||||||| || | || || || || |||     ||     ||     ||  |||
Db    211 ACGGCACCAACCTGACCATGGACGACGCGGCCCCCGCGCTCGCGGGGGACCTCA 270

Qy    394 CCGA 397
              || |
Db    271 CCCA 274
```

```
                        National Cancer Institute / NIH
                        Bldg. 31 Rm10A07 Bethesda, MD 20892
                        Email: cgapbs-r@mail.nih.gov
                        Tissue Procurement: Gerard Gradwohl (PNAS 97 P1607-1611,
                        2000)
                        cDNA Library Preparation: Catherine Lee, Endocrine Pancreas
                        Consortium
                        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                        DNA Sequencing by: Agencourt Bioscience Corporation
                        Clone distribution: MGC clone distribution information can be
                        found through the I.M.A.G.E. Consortium/LLNL at:
                        http://image.llnl.gov
                        Plate: IRBD19 row: h  column: 09
                        High quality sequence stop: 551.
FEATURES             Location/Qualifiers
     source          1..961
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6431780"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_137"
                     /note="Organ: pancreas; Vector: pSPORT1; Site 1: SalI;
                     Site 2: NotI; Library consists of a pool of clones
                     rearrayed from the following libraries: Melton normalized
                     rearrayed mouse pancreas 1 N1-MMS1, Amplified Melton mouse
                     mixed mouse pancreas 1 N1-MMS1, Amplified Melton mouse
                     islets 1 MIS1-A, and Kaestner ngn3 wt. Clones rearrayed in
                     the laboratory of K. Kaestner (University of
                     Pennsylvania). Note: this is a NIH_MGC Library."
ORIGIN

Query Match         13.2%;  Score 86.4;  DB 14;  Length 961;
Best Local Similarity  66.8%;  Pred. No. 1.8e-07;
Matches  123;  Conservative  0;  Mismatches  61;  Indels  0;  Gaps  0;

Qy    214 TGTGGTGTGAAAAGGGGCGTTCAGGTACTACTGACCACCATCGGCGCCCTTCGCGGCGTTTTG 273
              |||||||    |||   |||     |||||||| || ||    ||      | ||  ||
Db    281 TGCGAATGCGACCGGCGGGCCTGCAGGATGCTCGCTGACCACACAAGAGAGCCTTCGGCGCCGCCTTCT 340

Qy    274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGACTCTCATCTGCA 333
              ||||||| ||||| |||| |  || |||||||| ||||| |   || |   || |||||
Db    341 CGCTCATGGCCATCGGCATCGGCGACTACTGGCTGTACTCCAGCGGCCACATCTGCA 400

Qy    334 ACACCACCAACCTCACAGCAGGTGATGACGGACCACACCCATCGTGGGGGGGCCAGTGGCTCCT 393
              || ||||||||| || | || || || || ||    | ||     ||     ||  || |
Db    401 ACGGCACCAACCTGACCATGGACGACGCGGCCCCCGCGCTCGCGGGGGACCTCA 460

Qy    394 CCGA 397
              || |
Db    461 CCCA 464
```

RESULT 15
CF583776

```
LOCUS       CF583776      961 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  AGENCOURT_11364559 updated NIH_MGC_137 Mus musculus cDNA clone
            IMAGE:6431780 5', mRNA sequence.
ACCESSION   CF583776
VERSION     CF583776.1  GI:35197038
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 961)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
```

Search completed: May 22, 2004, 16:59:11
Job time : 2372 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: May 22, 2004, 15:39:32 ; Search time 312 Seconds
                                        (without alignments)
                                        9555.581 Million cell updates/sec

Title: US-10-060-066-2
Perfect score: 656
Sequence: 1 tatggctcactctgcacct.........tgttattttatttccttg 656

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications_NA:*
               1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
               2: /cgn2_6/ptodata/2/pubpna/US07_PCT_NEW_PUB.seq:*
               3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
               4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
               5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
               6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
               7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
               8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
               9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
               10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
               11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
               12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
               13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
               14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
               15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
               16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
               17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
               18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
               19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	170.8	26.0	1091	US-09-945-527-9	Sequence 9, Appli
2	96.4	14.7	1173	US-10-114-270-1	Sequence 1, Appli
3	96.4	14.7	1173	US-10-403-161-27	Sequence 27, Appl
4	86.4	13.2	984	US-09-883-483-1	Sequence 1, Appli
5	54.4	8.3	1252	US-10-276-774-825	Sequence 825, App
6	52	7.9	1383	US-10-276-774-780	Sequence 780, App
7	46.4	7.1	440	US-10-184-644-202	Sequence 202, App
8	46.4	7.1	440	US-10-184-63-202	Sequence 202, Appl
9	46.4	7.1	440	US-10-063-685-52	Sequence 52, Appl
10	45.8	7.0	956	US-10-027-632-31508	Sequence 31508, A
11	45.8	7.0	956	US-10-027-632-31508	Sequence 31508, A
12	44	6.7	497	US-10-424-599-133788	Sequence 133788,
13	43.8	6.7	791	US-10-424-599-64474	Sequence 64474, S
14	43.6	6.6	451	US-10-240-425-721	Sequence 721, App

c	15	43.2	6.6	71292	13	US-10-087-192-1942	Sequence 1942, Ap
c	16	43	6.6	805	13	US-10-424-599-61903	Sequence 61903, A
c	17	42.8	6.5	1926	15	US-10-294-804-3	Sequence 3, Appli
c	18	42.8	6.5	8705	15	US-10-291-230-14	Sequence 14, Appl
c	19	42.8	6.5	8705	15	US-10-291-249-14	Sequence 14, Appl
c	20	42.8	6.5	8705	17	US-10-273-678-16	Sequence 16, Appl
c	21	42.8	6.5	9600	15	US-10-279-751-1	Sequence 1, Appli
c	22	42.8	6.5	10233	15	US-10-050-898-283	Sequence 283, App
c	23	42.8	6.5	10285	15	US-10-050-902-283	Sequence 283, App
c	24	42.4	6.5	8837	15	US-10-207-655-56	Sequence 56, Appl
c	25	42.2	6.4	615	13	US-10-282-122A-11845	Sequence 11845, A
c	26	41.6	6.3	10343	13	US-10-655-847-11	Sequence 11, Appl
c	27	41.6	6.3	10343	13	US-10-160-807-11	Sequence 11, Appl
c	28	41.4	6.3	594	15	US-10-142-426-10	Sequence 10, Appl
c	29	41.4	6.3	594	15	US-10-123-155-10	Sequence 10, Appl
c	30	41.4	6.3	594	15	US-10-146-731-10	Sequence 10, Appl
c	31	41.4	6.3	594	15	US-10-140-472-10	Sequence 10, Appl
c	32	41.4	6.3	594	15	US-10-141-761-10	Sequence 10, Appl
c	33	41.4	6.3	594	15	US-10-141-885-10	Sequence 10, Appl
c	34	41.4	6.3	594	15	US-10-158-790-10	Sequence 10, Appl
c	35	41.4	6.3	594	16	US-10-137-871-10	Sequence 10, Appl
c	36	41.4	6.3	594	16	US-10-207-923-10	Sequence 10, Appl
c	37	41.4	6.3	594	16	US-10-141-756-10	Sequence 10, Appl
c	38	41.4	6.3	594	16	US-10-141-759-10	Sequence 10, Appl
c	39	41.4	6.3	594	16	US-10-141-805-10	Sequence 10, Appl
c	40	41.4	6.3	594	16	US-10-114-864-10	Sequence 10, Appl
c	41	41.4	6.3	12733	15	US-10-032-393-47	Sequence 47, Appli
c	42	41.4	6.3	12739	15	US-10-032-393-8	Sequence 8, Appli
c	43	41.4	6.3	114793	15	US-10-148-806-3	Sequence 3, Appli
c	44	41.2	6.3	528	13	US-10-424-599-42496	Sequence 42496, A
c	45	41.2	6.3	804	13	US-10-424-599-100535	Sequence 100535,

ALIGNMENTS

RESULT 1
US-09-945-527-9

Sequence 9, Application US/09945527

Publication No. US20030055588A1
GENERAL INFORMATION:
 APPLICANT: Robison, Keith E.
 TITLE OF INVENTION: No. US20030055588A1el Nucleic Acid Molecules Encoding
 TITLE OF INVENTION: Nucleic Acid and Protein Homologs
 FILE REFERENCE: 35800/237985
 CURRENT APPLICATION NUMBER: US/09/945,527
 CURRENT FILING DATE: 2001-08-29
 NUMBER OF SEQ ID NOS: 65
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 9
 LENGTH: 1091
 TYPE: DNA
 ORGANISM: Homo sapiens
US-09-945-527-9

Query Match   26.0%; Score 170.8; DB 10; Length 1091;
Best Local Similarity   78.8%;
Matches 219; Conservative   0; Mismatches 52; Indels   7; Gaps   1;

Qy   187 TGAAACGCTGGAATGAAGAGAGGGGTTTGTGTGTGTGAAAAGGGCGTTCAGGTACTACTGA 246
             ||||| |||||||||||||||||||||| |||||||| ||||||||||| ||||||| ||||
Db   31 TGAAGCGCTGGAAGACGGAGCGGGGCCTTGGTGCGGAGAGGGGGGTGCAGGTGCTGCTGA 90

Qy   247 CCACCATCCGGCGCGCCCTTCGCGGGGCTTTTGGCCTCATGACCATCAGCACTGACTACT 306
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91 CGACCGTGGCGCGCCCTTCGCGCGCCCTTCGGCCTCATGACCATCAGCGCCATCAGCTACT 150

Qy   307 GGCTCTACACAAGAGACTCTCATCTGCAACACCACCAAACCTCACAGC-------AGGTGAT 359
             |||| ||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   151 GGCTCTACACGCGCGCGCCCTCATCTGCAACACCACCAAACCTCACAGCCCCGGCGGGCGGGCC 210

Qy   360 GACGGACCACCCCATCGTGGGGGGCCAGTGGCTCCTCCGAGAAGAGACCCTGGGGGGCCTC 419
             ||||| || |||||| ||||| |||||||||||||||||||||||||||||| ||||||||
Db   211 GACGGACCACCCCATCGGGGGCGGGGCAGTGGCTCCTCCGAGAAGAGACCCTGGGGGGCCTC 210

```
Db          , 211 GGGACCCCTCACCTGCGGGGGGGCTGGGGGGGGCCTCGGAGAAGAAGGACCCCGCGCGGCCTC 270
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy          , 420 ACACATTCAGGCGCTTCTGGCGGATATGCTGCCTTGGGAAGG 457

Db            271 ACGCACTCCGGGCGCTTCTGGGAGGATTCTGCTGCCCTGGGAAGG 308
```

RESULT 2
US-10-114-270-1
; Sequence 1, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Tapier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US2004030110A1el
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: US/10/114,270
; PRIOR FILING DATE: 2002-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 1
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55)..(1024)

  Query Match   14.7%; Score 96.4; DB 13; Length 1173;
  Best Local Similarity   83.8%; Pred. No. 7.7e-19;
  Matches 109; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy   187 TGAAACGCTGGAATGAAGAGGAGGGGGGTTTTGTGGTGTGAAAAAGGCGGTTCAGGTACTCA 246
         || |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   38 TGAAGCGCTGGAACGAAATGCCGGGGCCTTCTGGTCGCAGAAGGGGGGTTGCAGGTACTGA 97

Qy   247 CCACCATCGCCGCCGCGCCTTCGCGGCGTTTTGGCCTCATGACCATCGCCATCAGCACTGACTACT 306
         || |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   98 CGACGGTGGGGCGGCGCCTTCGCGGCCCTCTGGTGCGGAGGACCATCGCCATCAGCACTGACTACT 157

Qy   307 GGCTCTACAC 316
         ||||||||||
Db   158 GGCTCTACAC 167
```

RESULT 3
US-10-403-161-27
; Sequence 27, Application US/10403161
; Publication No. US20040043930A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David et al.
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-573C
;  CURRENT APPLICATION NUMBER: US/10/403,161
;  CURRENT FILING DATE: 2003-03-31
;  PRIOR APPLICATION NUMBER: 60/370349
;  PRIOR FILING DATE: 2002-04-05
;  PRIOR APPLICATION NUMBER: 60/384543
;  PRIOR FILING DATE: 2002-05-30
;  PRIOR APPLICATION NUMBER: 60/370969
;  PRIOR FILING DATE: 2002-04-08
;  PRIOR APPLICATION NUMBER: 60/403748
;  PRIOR FILING DATE: 2002-08-15
;  PRIOR APPLICATION NUMBER: 60/372019
;  PRIOR FILING DATE: 2002-04-12
;  PRIOR APPLICATION NUMBER: 60/374379
;  PRIOR FILING DATE: 2002-04-22
;  PRIOR APPLICATION NUMBER: 09/779679
;  PRIOR FILING DATE: 2001-02-08
;  PRIOR APPLICATION NUMBER: 60/181045
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: 10/055877
;  PRIOR FILING DATE: 2002-01-22
;  PRIOR APPLICATION NUMBER: 60/262892
;  PRIOR FILING DATE: 2001-01-19
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: CuraSeqlist version 0.1
;  SEQ ID NO 27
;  LENGTH: 1173
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (55)..(1023)
US-10-403-161-27

  Query Match         14.7%;  Score 96.4;  DB 13;  Length 1173;
  Best Local Similarity   83.8%;  Pred. No. 7.7e-19;
  Matches 109;  Conservative  0;  Mismatches  21;  Indels  0;  Gaps  0;

Qy   187 TGAAACGCTGGAATGAAGAGGAGGGGGGTTTTGTGGTGTGAAAAAGGCGGTTCAGGTACTCA 246
         || |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    38 TGAAGCGCTGGAACGAAATGCCGGGGCCTTCTGGTCGCAGAAGGGGGGTTGCAGGTACTGA 97

Qy   247 CCACCATCGGCGCCGCTTCGCGGCGCTTTTGGCCTCATGACCATCAGCACACTGACTACT 306
         || |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    38 TGAAGCGCTGGAACGAAATGCCGGGGCCTTCTGGT... 97
```

```
Db   98  CGACGGTGGGGCGCGCCTTCGCCGCCTTCATGACCATCGCCATCAGCACACTGACTACT  157
Qy  307  GGCTCTACAC  316
Db  158  GGCTCTACAC  167
```

RESULT 4
US-09-883-483-1
; Sequence 1, Application US/09883483
; Patent No. US20020128454A1
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, David Malcolm
; APPLICANT: HAYES, Philip David
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30180-D1
; CURRENT APPLICATION NUMBER: US/09/883,483
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: UK 9819592.8
; PRIOR APPLICATION FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: 09/392,014
; PRIOR APPLICATION FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 984
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-883-483-1

Query Match   13.2%; Score 86.4; DB 9; Length 984;
Best Local Similarity   66.8%; Pred. No. 8e-16;
Matches 123; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```
Qy  214  TGTGGTGTGAAAAGGGCGTTCAGGTAGTACTACTGACCACCATCGGGCGGCGGCCTTCGCTTTTTGGCCT  273
Db    5  TCGGATGCGACACCCGGGGCTGCAGAATGCTGCTGACCACCACGGACGCCTTCGCGGCCTTCT  64
Qy  274  GCCTCATCGACCATCAGCACTGACTACTGCCTCTACACAAAGAGCTTCATCTGCA  333
Db   65  CGGCTCATGGCCATCGCCATCGGCACCTACTGGCTGTACTCCAGCGGCACATCTGCA  124
Qy  334  ACACCACCAAACCTTCACACAGCAGGGTGATGACGGACACACCCCATCGTGTGGGGGGCCAGTGGCTCCT  393
Db  125  ACGGCACCAAACCTGACCATGGACGACGAGGGCCCCCGCGCCGCCCGGGCCACCTCA  184
Qy  394  CCGA  397
Db  185  CCCA  188
```

RESULT 5
US-10-276-774-825
; Sequence 825, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR APPLICATION FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR APPLICATION FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 825
; LENGTH: 1252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-825

Query Match   8.3%; Score 54.4; DB 13; Length 1252;
Best Local Similarity 70.2%; Pred. No. 4.7e-06;
Matches 73; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy  218  GTGTGAAAAGGGGCGGTTCAGGTAGTACTACTGACCACCATCGGGCGGCGGCCTTCGCGGCCTTCGGCCT  277
Db  291  GTTTGATCGAGGGTGTTCAAATGCTTTTAACCACCGTTGGTTGCTTCGGCTCGGCTTCCTTCAGCCT  350
Qy  278  CATGACCATCGCCATCAGCACAGCTGACTACTGGCTCTACACAAGAG  321
Db  351  GATGACCATAGCCTGTGGGAACCGACTATTCGGCTCTACTCCAGAG  394
```

RESULT 6
US-10-276-774-780
; Sequence 780, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR APPLICATION FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR APPLICATION FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 780
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-780

Query Match   7.9%; Score 52; DB 13; Length 1383;
Best Local Similarity 65.5%; Pred. No. 2.6e-05;
Matches 76; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```
Qy  218  GTGTGAAAAGGGGCGTTCAGGTAGTACTACTGACCACCATCGGGCGGCGGCCTTCGCGGCCTTTGCGCCT  277
Db  429  GTGTGACAGAGGTATCCAGATAGTTGATCACCACTGTGAGGACGCCTTTGCCCGGCTTTTAGTTT  488
Qy  278  CATGACCATCGCCATCAGCACTGACTACTCGGCTCTACACAAGAGACTCTCATCTCTGCA  333
Db  489  AATGACCATTGCCAGTGGGCACGGACTACTCGGTTATATTCCAGAGAGGTGTGTGTCCAGGA  544
```

RESULT 7
US-10-184-644-202/c
; Sequence 202, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 202
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-202

Query Match   7.1%; Score 46.4; DB 15; Length 440;
Best Local Similarity   22.0%; Pred. No. 0.00091;
Matches 38; Conservative 58; Mismatches 77; Indels 0; Gaps 0;

Qy   5 GCTTCACTCTGCACCTTCGCCTCTGCAGCCACCTGCTGCTGCCCCCGGGCCCCCTCTGGC 64
         :||:.|| ||||||||: |||: |||||||||||||||||||||: :||:.|||||||||
Db 388 SCWSSC.CYRNHCSCC..YNC.MS..YXNSSBC.CY.C..CS.CSCYT.NC.M.GC.MDC 329

Qy   65 CCCTTGAGGCCCCCCCCAGCTTCTTGCCTGTCGTGTGATCCCCCAGCCGCCGGGCACGGCCCC 124
         :||:.| |||||||: |||: |||: |||: |||||||||: :||:.|||||||||: :||:.
Db 328 NCCCSCCDNCSSSCASSCWSS.SCSHCYSCCSNCSSCCSSSCCSSSCSSCSSCCSSSSCCSSSCCS 269

Qy 125 GCCTCCGCTGCCCCGGTGGTGGGCCCACGGGCCCCCCCGGTTGCCAGTGGTGAAAC 177
         :||: :|||:.|||||||||||||||||||||: :||:.|||||||||: :||:.|
Db 268 SSCNNNHCNSCSCSSCS.SCSCCCSNSSCCCSCS...NAGC.N.NSTYBSCRC 216

RESULT 8
US-10-184-634-202/c
; Sequence 202, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 612
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 202
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-202

Query Match   7.1%; Score 46.4; DB 15; Length 440;
Best Local Similarity   22.0%; Pred. No. 0.00091;
Matches 38; Conservative 58; Mismatches 77; Indels 0; Gaps 0;

Qy   5 GCTTCACTCTGCACCTTCGCCTCTGCAGCCACCTGCTGCTGCCCCCGGGCCCCCTCTGGC 64
         :||:.|| ||||||||: |||: |||||||||||||||||||||: :||:.|||||||||
Db 388 SCWSSC.CYRNHCSCC..YNC.MS..YXNSSBC.CY.C..CS.CSCYT.NC.M.GC.MDC 329

Qy   65 CCCTTGAGGCCCCCCCCAGCTTCTTGCCTGTCGTGTGATCCCCCAGCCGCCGGGCACGGCCCC 124
         :||:.| |||||||: |||: |||: |||: |||||||||: :||:.|||||||||: :||:.
Db 328 NCCCSCCDNCSSSCASSCWSS.SCSHCYSCCSNCSSCCSSSCCSSSCSSCSSCCSSSSCCSSSCCS 269

Qy 125 GCCTCCGCTGCCCCGGTGGTGGGCCCACGGGCCCCCCCGGTTGCCAGTGGTGAAAC 177
         :||: :|||:.|||||||||||||||||||||: :||:.|||||||||: :||:.|
Db 268 SSCNNNHCNSCSCSSCS.SCSCCCSNSSCCCSCS...NAGC.N.NSTYBSCRC 216

; GENERAL INFORMATION:
; APPLICANT: Eaton,Dan L.
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,Christopher J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,685
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 52
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-685-52

Query Match   7.1%; Score 46.4; DB 15; Length 440;
Best Local Similarity   22.0%; Pred. No. 0.00091;
Matches 38; Conservative 58; Mismatches 77; Indels 0; Gaps 0;

Qy   5 GCTTCACTCTGCACCTTCGCCTCTGCAGCCACCTGCTGCTGCCCCCGGGCCCCCTCTGGC 64
         :||:.|| ||||||||: |||: |||||||||||||||||||||: :||:.|||||||||
Db 388 SCWSSC.CYRNHCSCC..YNC.MS..YXNSSBC.CY.C..CS.CSCYT.NC.M.GC.MDC 329

Qy   65 CCCTTGAGGCCCCCCCCAGCTTCTTGCCTGTCGTGTGATCCCCCAGCCGCCGGGCACGGCCCC 124
         :||:.| |||||||: |||: |||: |||: |||||||||: :||:.|||||||||: :||:.
Db 328 NCCCSCCDNCSSSCASSCWSS.SCSHCYSCCSNCSSCCSSSCCSSSCSSCSSCCSSSSCCSSSCCS 269

Qy 125 GCCTCCGCTGCCCCGGTGGTGGGCCCACGGGCCCCCCCGGTTGCCAGTGGTGAAAC 177
         :||: :|||:.|||||||||||||||||||||: :||:.|||||||||: :||:.|
Db 268 SSCNNNHCNSCSCSSCS.SCSCCCSNSSCCCSCS...NAGC.N.NSTYBSCRC 216

RESULT 10
US-10-027-632-31508
; Sequence 31508, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US/10/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 31508
; LENGTH: 956
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

RESULT 9
US-10-063-685-52/c
; Sequence 52, Application US/10063685
; Publication No. US20030180909A1

; LOCATION: (1)...(956)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-31508

Query Match   7.0%; Score 45.8; DB 13; Length 956;
Best Local Similarity   55.3%;
Matches 83; Conservative   0; Mismatches 67; Indels   0; Gaps   0;

```
Qy     9 CACTCTGCAACCTTCGCCTCTGCAGCCACCTGCTGCTGCGCCCCCGGCCCCTCTGGCCCCT 68
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   392 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 451

Qy    69 TGAGGCCCCCCAGCTTCTGCCTGTGTGATCCCCNCNCCAGCCGCCCGGCACGGCCCCCCT 128
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   452 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCGGNCCCGGGCCCCCCGGN 511

Qy   129 CCGCTGCCCGGGTGGGTGGCCCACGGCCCCC 158
           |||||||||||||||||||||||||||||||
Db   512 GCGGCGACACCCGGCGGGGGCCCCCTGCCCC 541
```

RESULT 11
US-10-027-632-31508
; Sequence 31508, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphism in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31508
; LENGTH: 956
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(956)
; OTHER INFORMATION: n = A,T,C or G

RESULT 12
US-10-424-599-133788
; Sequence 133788, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 133788
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(497)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91816C.1
US-10-424-599-133788

Query Match   6.7%; Score 44; DB 13; Length 497;
Best Local Similarity   55.1%;
Matches 86; Conservative   0; Mismatches 70; Indels   0; Gaps   0;

```
Qy     4 GGCTTCACTCTGCAACCTTCGCCTCTGCAGCCACCTGCTGCTGCGCCCCCGCCCCTCTGG 63
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309 GCCCCCCCCCTGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 368

Qy    64 CCCCTTGAGGCCCCCCAGCTTCTGCCTGTGTGATCCCCCCCCCAGCCGCCGGCACGGGCCC 123
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   369 GCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCTCCCCCCCCCCCCCCCCCCCCCCCCCC 428

Qy   124 CGCCTCCGCTGCCCCGGTGGTGGGCCCACGGCCCCC 159
           ||||||||||||||||||||||||||||||||||||
Db   429 CCCCCCCCCCCCCCCCCCCCTGCCCCCCCCCCCCCCC 464
```

RESULT 13
US-10-424-599-64474/c
; Sequence 64474, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 64474
; LENGTH: 791
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29231C.1
US-10-424-599-64474

Query Match   6.7%; Score 43.8; DB 13; Length 791;
Best Local Similarity   54.7%;
Matches 87; Conservative   0; Mismatches 72; Indels   0; Gaps   0;

```
Qy     9 CACTCTGCAACCTTCGCCTCTGCAGCCACCTGCTGCTGCGCCCCCGGCCCCTCTGGCCCCT 68
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   392 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 451

Qy    69 TGAGGCCCCCCAGCTTCTGCCTGTGTGATCCCNCNCCACCCGCCCGGCACGGCCCCCCT 128
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   452 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCGGNCCCGGGCCCCCCGGNGGCCA 511

Qy   129 CCGCTGCCCGGGTGGGTGGCCCACGGCCCCC 158
           |||||||||||||||||||||||||||||||
Db   512 GCGCGCGACACCCGGCGGGGGCCCCCTGCCCC 541
```

```
Qy    8 TCACTCTTGCACTTTCGCCTTCTTGCAGCCACTGCTTCGCTGCCCCCTCTGGCCCC 67
        ||||||||||||||||||||||||||||||||||||||||||||||||   ||||
Db  440 TCTCTCGCCCCCCCAGCCACTGCTTCGCTCGCCCCCCCCGCCCCCCCCCCC 381

Qy   68 TTGAGGCCCCCCCAGCTTCTGCCTGTGCTTGTGTGATCCCCCCCCCCCCCCCCCCCCG 127
        ||||||||||||||||||   ||||||||||||||||   |||||||||||
Db  380 CTCCCCCCCTCCCCCCCTTCTCCCCCCAGCCGCCGGCCACGGCCCCCCGCCC 321

Qy  128 TCGGCTGCCCCGGGGTGGGGCCCCACGGCCCCCCCCCCCCCCCCGGTTGCC 166
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  320 CCTCCTCCCGCCGGTCCACCCCCCCCCCCCCCCCCTGCCCCC 282
```

RESULT 14
US-10-240-425-721/c
; Sequence 721, Application US/10240425
; Publication No. US20040035502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 721
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040035502A1 AI640222
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(451)
; OTHER INFORMATION: n = a or c or g or t
US-10-240-425-721

Query Match   6.6%; Score 43.6; DB 13; Length 451;
Best Local Similarity 49.4%; Pred. No. 0.0065;
Matches   76; Conservative   0; Mismatches 78; Indels   0; Gaps   0;

```
Qy    6 CTTCACTCTTGCACTTTCGCCTTCTTGCAGCCACTGCTTCGCTGCCCCCTCTCGGCC 65
        ||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db  418 CCNCNCCCCCNCCCCCCNCCCCCCNCCCCCCNNCCCCCCCCCCCCCCCCCCC 359

Qy   66 CCTTGAGGCCCCCCCAGCTTCTGCCTGTGCTTGTGTGATCCGCGGGCACGGCCCCG 125
        ||||||||||||||||   |||||||||||||||||||   |||||||||||
Db  358 CCCCCNNNCNCCCCCNCCCCCCNNNCCCCCCCCCCCCCCNNCCCCCCCC 299

Qy  126 CCTCCCGCCTGCCCCGGGGTGGGGCCCCACGGCCCCCC 159
        ||||||||||||||||||||||||||||||||||||||||
Db  298 CCCCCCCCCCCCNCCCCCCCNCCCCCCCC 265
```

RESULT 15
US-10-087-192-1942/c
; Sequence 1942, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER

; FILE REFERENCE: 52945200012Z
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
; LENGTH: 71292
US-10-087-192-1942

Query Match   6.6%; Score 43.2; DB 13; Length 71292;
Best Local Similarity 55.3%; Pred. No. 0.046;
Matches   84; Conservative   0; Mismatches 68; Indels   0; Gaps   0;

```
Qy   15 GCACCTTCGCCTCTGCAGCCACCTGCTTCGCTCGCTGCCCCCGGCCCCCTCTGGCCCTTGAGGC 74
        ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db 9654 GCCGCCTTCCGCCTGCGGACACGCCCTCACTTCCCCCCCCCCCCCCCCCCCCCCC 9595

Qy   75 CCCCCAGCTTCTGCCTGTGCTGTGATCCCCCCCAGCCCACGGGCCCCCGCCCTTCCGCTG 134
        |||||||||||||||   ||||||||||||||||   ||||||||||||||||||
Db 9594 CCGCCTGCCCCTGGCCCACGCCCTGGGCCCACGCCCGCGTCCGCGCGGGGCGCGC 9535

Qy  135 CCCCGGGTGGTGGGCCCACGGCCCCCCCCCGGTTGCC 166
        |||||||||||||||||||||||||||||||||||||||
Db 9534 CCTCCAGGACACCCGCCCGCGGGCCCCCAGGGTCCC 9503
```

Search completed: May 22, 2004, 17:06:08
Job time : 314 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004, Compugen Ltd.

Run on: May 22, 2004, 15:34:27 ; Search time 66 Seconds
(without alignments)
5515.876 Million cell updates/sec

Title: US-10-060-066-2
Perfect score: 656
Sequence: 1 tatggcttcacctctgcacct.........tgtctattcttattcttcttg 656

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	DB ID	Description
1	86.4	13.2	984	3	US-09-392-014-1	Sequence 1, Appl
2	83.2	12.7	1071	4	US-09-123-030-11	Sequence 11, Appl
3	53.6	8.2	1477	4	US-09-123-030-9	Sequence 9, Appli
4	52.8	8.0	1558	4	US-09-123-030-7	Sequence 7, Appli
5	45.8	7.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
6	45.8	7.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
7	42.8	6.5	1926	4	US-09-249-585A-2	Sequence 2, Appl
8	42.8	6.5	1926	4	US-09-410-399-3	Sequence 3, Appl
9	42.8	6.5	2580	3	US-09-050-863-2	Sequence 2, Appl
10	42.8	6.5	2580	3	US-09-359-081-2	Sequence 2, Appl
11	42.8	6.5	5452	2	US-09-130-114-1	Sequence 1, Appl
12	42.8	6.5	8705	4	US-09-647-344A-14	Sequence 14, Appl
13	42.8	6.5	9600	4	US-09-910-647-1	Sequence 1, Appl
14	42.8	6.5	9600	4	US-09-620-925-1	Sequence 1, Appl
15	42.8	6.5	10596	1	US-07-884-811-15	Sequence 15, Appl
16	42.8	6.5	10596	1	US-07-885-971-15	Sequence 15, Appl
17	42.8	6.5	10596	1	US-08-087-783A-15	Sequence 15, Appl
18	42.8	6.5	10596	1	US-08-194-087-15	Sequence 15, Appl
19	42.8	6.5	10596	5	US-08-194-087B-15	Sequence 15, Appl
20	42.8	6.5	16080	4	PCT-US93-04648-15	Sequence 15, Appl
21	42.8	6.5	16080	4	US-09-724-566A-48	Sequence 48, Appl
22	41.4	6.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
23	41.4	6.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
24	41.2	6.3	320	3	US-09-165-264-7	Sequence 7, Appli
25	39.8	6.1	1347	4	US-09-140-804-1	Sequence 1, Appli
26	39.8	6.1	1347	4	US-09-686-838B-1	Sequence 1, Appl
27	39.8	6.1	1377	4	US-09-866-028-41	Sequence 41, Appl
c 28	39.6	6.0	320	3	US-09-165-264-11	Sequence 11, Appl
c 29	39.6	6.0	1382	2	US-09-016-434-1256	Sequence 1256, Ap
c 30	39.6	6.0	152331	3	US-09-128-155-16	Sequence 16, Appl
c 31	39.4	6.0	4257	2	US-08-690-473-1	Sequence 1, Appli
c 32	39.4	6.0	4257	3	US-09-259-821A-1	Sequence 1, Appli
c 33	39.4	6.0	4257	3	US-08-843-659-1	Sequence 1, Appl
c 34	39.4	6.0	12001	1	US-08-458-568A-11	Sequence 11, Appl
c 35	38.2	5.8	12001	1	US-09-165-264-8	Sequence 8, Appl
c 36	38.2	5.8	319	3	US-09-165-264-14	Sequence 14, Appl
c 37	38	5.8	320	3	US-08-232-463-14	Sequence 14, Appl
c 38	38	5.8	7218	1	US-09-165-264-12	Sequence 12, Appl
c 39	37.8	5.8	318	3	US-09-165-264-13	Sequence 13, Appl
c 40	37.8	5.8	320	1	US-09-621-976-17202	Sequence 17202, A
c 41	37.6	5.7	12001	2	US-09-016-434-1286	Sequence 1286, Ap
c 42	37.2	5.7	364	4	US-09-165-264-13	Sequence 13, Appl
c 43	37.2	5.7	1062	4	US-08-134-231C-33	Sequence 33, Appl
c 44	37.2	5.7	1062	4	US-09-023-655-1327	Sequence 1327, Ap
c 45	37.2	5.7	1062	4	US-08-728-160-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-392-014-1
; Sequence 1, Application US/09392014
; Patent No. 6274380
; GENERAL INFORMATION:
; APPLICANT: David Malcolm Duckworth
; APPLICANT: Philip David Hayes
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30180
; CURRENT APPLICATION NUMBER: US/09/392,014
; CURRENT FILING DATE: 1999-09-08
; EARLIER APPLICATION NUMBER: UK 9819592.8
; EARLIER FILING DATE: 1998-09-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 984
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-392-014-1

Query Match 13.2%; Score 86.4; DB 3; Length 984;
Best Local Similarity 66.4%;
Matches 123; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```
Qy   214 TGTGGTGTGAAAAAGGCGGTTCAGGTTACTACTGACCACCATCGGGCCGCCCTTTGCGCGGCTTTTG 273
             |||| |||     |||| |||   |  || |||||||||||  |||  ||| ||||||   ||||||||
Db     5 TGCGGATNGCGACCGCGGGCTGCCAGANTCGTGCTGACCACAGGGCGGAGCCTTCGCCGCCTTCT 64

Qy   274 GCCTCCATGACCAATCGCCATAGCACTGACTTACTGGCTTCTACACAAGAGCTTCTCATCTGCA 333
             |||||||| ||| ||| |||||| ||| |  |||||||| || ||   |||| |||| ||||||
Db    65 CGCTCCATGGCCCATCCGCCATCGGCACGGCACCGACTACTGGCTGTGACTACTGGCTGTGCA 124

Qy   334 ACACCACCAACCTCACAGCAGGTGATGACCGACCACCCCATCGTGGGGGCAGTGGCTCCT 393
             || ||||| |||| ||  |  |  ||| ||| |||| || |||  || || ||  |  |
Db   125 ACGGCACCAACCTGACCATGGACGACGGCGCCCCGCCCGCGCGGCGCCCGACCTCA 184

Qy   394 CCGA 397
             ||||
Db   185 CCCA 188
```

RESULT 2
US-09-123-030-11
; Sequence 11, Application US/09123030
; Patent No. 6365337
; GENERAL INFORMATION:
; APPLICANT: Lette, Verity A.
; APPLICANT: Frankel, Wayne N.
; APPLICANT: Campbell, Kevin P.

; APPLICANT: Felix, Ricardo
; APPLICANT: Biddlecome, Gloria
; TITLE OF INVENTION: Genes Encoding Neuronal Voltage-Gated Calcium Channel
; TITLE OF INVENTION: Gamma Subunits
; FILE REFERENCE: US App. 09/123,030
; CURRENT APPLICATION NUMBER: US/09/123,030
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 11
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(21)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22)..(1002)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1003)..(1071)
US-09-123-030-11

Query Match   12.7%;   Score 83.2;   DB 4;   Length 1071;
Best Local Similarity   65.8%;   Pred. No. 4.1e-13;
Matches 121;   Conservative 0;   Mismatches 63;   Indels 0;   Gaps 0;

Qy   214 TGTGGTGTGAAAAGGGCGGTTCAGGTACTACTGACCACCATCGGCGCGCCTTCGCGGCTTTTG 273
Db   26 TGCGATCGGACCGCGGCGTGCAGATGCTGCTGACCACCACGGCGGAGCCCTCGCCGCCTTCT   85

Qy   274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTTCTTACACAAGAGCTCTCATCTGCA 333
Db   86 CGCTCATGGCCATCGCCATCGGCCACCGACTACCGGCTTGTACTCCAGCGCGCACATCTGCA 145

Qy   334 ACACCACCAACCTCACAGACAGGTTGATGACGGACCACCCCCATCGTGTGGGGGGGCAGTGGCTCCT 393
Db   146 ACGGCACCAAACTGACCATGACGACGCGGCCCCCGCCGCCGCTCGGCGGCGACCTCA 205

Qy   394 CCGA 397
Db   206 CCCA 209

RESULT 3
US-09-123-030-9
; Sequence 9, Application US/09123030
; Patent No. 6365337
; GENERAL INFORMATION:
; APPLICANT: Letts, Verity A.
; APPLICANT: Frankel, Wayne N.
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Felix, Ricardo
; APPLICANT: Biddlecome, Gloria
; TITLE OF INVENTION: Genes Encoding Neuronal Voltage-Gated Calcium Channel
; TITLE OF INVENTION: Gamma Subunits
; FILE REFERENCE: US App. 09/123,030
; CURRENT APPLICATION NUMBER: US/09/123,030
; FILE FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 9
; LENGTH: 1477
; TYPE: DNA
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(487)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (488)..(1432)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1433)..(1477)
US-09-123-030-9

Query Match   8.2%;   Score 53.6;   DB 4;   Length 1477;
Best Local Similarity   66.4%;   Pred. No. 4.6e-05;
Matches 77;   Conservative 0;   Mismatches 39;   Indels 0;   Gaps 0;

Qy   218 GTGTGAAAAGGGCGGTTCAGGTACTACTGACCACCATCGGCGCGCCTTCGCGGCTTTTGGGCCT 277
Db   496 GTGTGACACAGAGGTATCCAGAAGTTGATCACTACTACTGTAGGAGCCTTCGCAGCTTTAGTTT 555

Qy   278 CATGACCATCGCCATCAGCACTGACTACTGGCTCTACAACAAGAGCTCTCATCTGCA 333
Db   556 AATGACCATTGCAGTGGGCACCGGACTACTGGCTATATTCCAGGAGGTGTGTGCAGGA 611

RESULT 4
US-09-123-030-7
; Sequence 7, Application US/09123030
; Patent No. 6365337
; GENERAL INFORMATION:
; APPLICANT: Letts, Verity A.
; APPLICANT: Frankel, Wayne N.
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Felix, Ricardo
; APPLICANT: Biddlecome, Gloria
; TITLE OF INVENTION: Genes Encoding Neuronal Voltage-Gated Calcium Channel
; TITLE OF INVENTION: Gamma Subunits
; FILE REFERENCE: US App. 09/123,030
; CURRENT APPLICATION NUMBER: US/09/123,030
; FILE FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 7
; LENGTH: 1558
; TYPE: DNA
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(389)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (390)..(1361)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1361)..(1558)
US-09-123-030-7

Query Match   8.0%;   Score 52.8;   DB 4;   Length 1558;
Best Local Similarity   69.2%;   Pred. No. 7.7e-05;
Matches 72;   Conservative 0;   Mismatches 32;   Indels 0;   Gaps 0;

Qy   218 GTGTGAAAAGGGCGGTTCAGGTACTACTGACCACCATCGGCGCGCCTTCGCGGCTTTTGGGCCT 277
Db   398 GTTTGATCGAGGTGTTCAAATGCTTTAACCACCGTTGGTGCTTTCGCTGCCTTCAGCTT 457

Qy   278 CATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAG 321
Db   458 GATGACCATCGCTGTGGGAACCGACTATTGGCTGTACTCCAGAG 501

RESULT 5
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION:
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match   7.0%; Score 45.8; DB 3; Length 4403765;
Best Local Similarity 54.4%; Pred. No. 0.12;
Matches 92; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy   5 GCTTCACTCTGCACCTTCGCCTTCTGCAGGCCACTGGCTGCCCCCGCGCCCCCTCTGGC 64
            |||||||||||||||||||| |||| ||||||||||| || |||||||||||| ||||||
Db 1631498 GCTTCGCGCCGCGCGCGGCGGCGGAAGGCGCCTTCGCCGCCTGGCCGGGTTGGCGCTGGC 1631557

Qy   65 CCCTTGAGGCCCCCGCGGGAGCACCCCCAGCTTCTGCCTGTTGATCCCCCAGCCTGCCCC 124
            ||||| ||||||||||| ||||||| ||||||||||| ||||| ||||||||| |||||||
Db 1631558 CCCGTCACGGGAGCACCCCCGCGCCGCGCCCCACGGCTCACGCCCGCGGCGGCGTCGGT 1631617

Qy   125 GCCTCCGCTGCCCCGGCGTGGTGGCCCCACGGCGCCCCCGGGTTGCCAGTGGTG 173
            |||||||| |||| ||||| |||||||||||||||||||||||||| |||||||||
Db 1631618 GACTACGGCCTTGACGGCCCGCGGCGGGCCCGGTTGCGCGCGGGG 1631666

RESULT 6
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match   7.0%; Score 45.8; DB 3; Length 4411529;
Best Local Similarity 54.4%; Pred. No. 0.12;
Matches 92; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy   5 GCTTCACTCTGCACCTTCGCCTTCTGCAGCCACTGGCTGCCCCCGCGCCCCCTCTGGC 64
            |||||||||||||||||||| |||| ||||||||||| || |||||||||||| ||||||
Db 1631638 GCTTCGCGCCGCGCGCGGCGGCGGAAGGCGCCTTCGCCGCCTGGCCGGGTTGGCGCTGGC 1631697

Qy   65 CCCTTGAGGCCCCCCAGCTTCTGCTGTGCTGTGATCCCCCACGCGCCGGCAGGGCCCC 124
            ||||| ||||||||||| |||||| || ||||| ||||||||| || |||||||||||||
Db 1631698 CCCGTCACGGGGAGACACCCCCGGCGCCCCACGGCCGCCGCGCCGGCT 1631757

Qy   125 GCCTCCGCTGCCCCGGTGGTGGGCCCCACGGCGCCCCCGGGTTGCCAGTGGTG 173
            |||||||| |||| |||||||||||||||||||||||||||||| |||||||||
Db 1631758 GACTACGGCCTTGACGGCCGCGGCGGGCCCGGTTGCGCGCGGGCG 1631806

RESULT 7
US-09-249-585A-2/c
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA

US-09-249-585A-2

Query Match   6.5%; Score 42.8; DB 4; Length 1926;
Best Local Similarity 58.7%; Pred. No. 0.042;
Matches 74; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy   13 CTGCACCTTCGCCTTCTGCAGCCACTGGCTGCCCCCGCGCCCCCTCTGGCCCCTTGAG 72
            ||||||| |||| ||||||||| || ||||| ||||||||||| |||||||||||||||||
Db   982 CTCCACCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTCCCCCCTCCTG 923

Qy   73 GCCCCCCAGCTTCTGCCTGTTGTGCTGTGAATCCCCCCAGCCGCGGGCACGGGCCCCGC 132
            ||||||||||||| ||||| || ||||||||||||||||||| ||||||||||||| ||||
Db   922 CCCCTTCCTGCTCCTGCCCCTCCCTGCTCCCTCCTGCTCCCGGC 863

Qy   133 TGCCCC 138
            ||||||
Db   862 CTGCCC 857

RESULT 8
US-09-410-399-3/c
; Sequence 3, Application US/0910399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Cotter, Murray A.
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
; TITLE OF INVENTION: to Genomic Host DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 3
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein-Barr virus

US-09-410-399-3

Query Match   6.5%; Score 42.8; DB 4; Length 1926;
Best Local Similarity 58.7%; Pred. No. 0.042;
Matches 74; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy   13 CTGCACCTTGCCCTCTGCAGCCACCTGCTCGCTGCGCCCCGGCCCCCTCTGGCCCCTTGAG 72
            ||||||| |||| ||||||||| || |||||||||||||||||| ||||||||||||||||
Db   982 CTCCACCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTCCCCCCTCCTG 923

Qy   73 GCCCCCCAGCTTCTGCCTGTTGACTGTGATCCCCCCAGCCGCGACGGGCCCCGCCTCCGC 132
            ||| |||||||| ||||| || |||| || |||||| ||||||||||||||||||| ||||
Db   922 CCCCTTCCTGCTCCTCCCTGCTCCCCCTCCTGCTCCCGGGCGGGCCCCGGC 863

Qy   133 TGCCCC 138
            ||||||
Db   133 TGCCCC 138

Db   862 CTGCCC 857
              |||

RESULT 9
US-09-050-863-2/c
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
;   APPLICANT: Lao, Ying
;   APPLICANT: Hiang, Betty
;   APPLICANT: Payan, Don
;   TITLE OF INVENTION: Mammalian Protein Interaction Cloning
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;   STREET: 4 Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94111-4187
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/050,863
;   FILING DATE: 30-MAR-1998
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Silva, Robin M.
;   REGISTRATION NUMBER: 38,304
;   REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 781-1989
;   TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2580 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA
US-09-050-863-2

Query Match   6.5%;   Score 42.8;   DB 3;   Length 2580;
Best Local Similarity 58.7%;   Pred. No. 0.046;
Matches   74;   Conservative   0;   Mismatches 52;   Indels   0;   Gaps   0;

Qy   13 CTGCACCTTCGCCTCTCGCAGCCACCTGCTTCGCTTGCCCCGGCGCCCCCTCTGGGCCCTTGAG 72
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1365 CTCCAACTTCCTGCTTCCTGCTTCGGTGTTGATCCCCGGCGCACGGCGCCCCCTCTGGGCCCTTGAG 1306

Qy   73 GCCCCCCAGCTTCTGCCTGTGCTTGTGATCCCCCGGCGCACGGCGCCCCGCCCTCCGC 132
                |||||||||||| |||||||||||||||| |||||||||||||||||||||||||
Db   1305 CCCCTCCTGCTTCCTGCCCCTTCCTGCTCTCGCCCCTCCTCGCCCCTCCTCGCC 1246

Qy   133 TGCCCC 138
                ||||||
Db   1245 CTGCCC 1240

RESULT 10
US-09-359-081-2/c
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
;   APPLICANT: Lao, Ying
;   APPLICANT: Hiang, Betty
;   APPLICANT: Payan, Don
;   TITLE OF INVENTION: Mammalian Protein Interaction Cloning

---

; NUMBER OF SEQUENCES: 5   System
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;   STREET: 4 Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94111-4187
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/359,081
;   FILING DATE: 22-Jul-1999
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 09/050,863
;   FILING DATE: <Unknown>
;   CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Silva, Robin M.
;   REGISTRATION NUMBER: 38,304
;   REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 781-1989
;   TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2580 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

Query Match   6.5%;   Score 42.8;   DB 4;   Length 2580;
Best Local Similarity 58.7%;   Pred. No. 0.046;
Matches   74;   Conservative   0;   Mismatches 52;   Indels   0;   Gaps   0;

Qy   13 CTGCACCTTCGCCTCTCGCAGCCACCTGCTTCGCTTGCCCCGGCGCCCCCTCTGGGCCCTTGAG 72
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1365 CTCCAACTTCCTGCTTCCTGCTTCGGTGTTGATCCCCGGCGCACGGCGCCCCCTCTGGGCCCTTGAG 1306

Qy   73 GCCCCCCAGCTTCTGCCTGTGCTTGTGATCCCCCGGCGCACGGCGCCCCGCCCTCCGC 132
                |||||||||||| |||||||||||||||| |||||||||||||||||||||||||
Db   1305 CCCCTCCTGCTTCCTGCCCCTTCCTGCTCTCGCCCCTCCTCGCCCCTCCTCGCC 1246

Qy   133 TGCCCC 138
                ||||||
Db   1245 CTGCCC 1240

RESULT 11
US-09-130-114-1
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
;   APPLICANT: Horlick, Robert A.
;   APPLICANT: Damaj, Bassam B.
;   APPLICANT: Robbins, Alan K.
;   TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
;   TITLE OF INVENTION: From Multiple Transfected Episomes
;   FILE REFERENCE: 0867/1D903US1
;   CURRENT APPLICATION NUMBER: US/09/130,114
;   CURRENT FILING DATE: 1998-08-06
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 5452
;   TYPE: DNA

; ORGANISM: VERNA
US-09-130-114-1

Query Match   6.5%;   Score 42.8;   DB 2;   Length 5452;
Best Local Similarity 58.7%;   Pred. No. 0.062;
Matches 74;   Conservative 0;   Mismatches 52;   Indels 0;   Gaps 0;

Qy   13 CTGCACCTTGGCCTCTGCAGCCACCTGCTGCCCCGGCCCCCTCTGGCCCCTTGAG 72
         ||||||||||| ||| |||||||||||| |||||||||||||||| ||||||||
Db 1440 CTCCACCTCCTGTCCTCCTGCCCCTCCTGCGGCCCCTCCTGCCCCTCCTG 1499

Qy   73 GCCCCCCAGCTTCTGCCTGTGCTGTGATCCCCCAGCCAGCGGGGCACGGCCCCGGCCTTCCGC 132
         |||||| ||| ||||||||| || |||||||||| ||||||||||||||||| |||||||
Db 1500 CCCTTCCTGCTCCTGCCCCTCCTGCCCCCTCGTCCCCTGCTCCCCTTCCTC 1559

Qy   133 TGCCCC 138
         ||||||
Db 1560 CTGCCC 1565

RESULT 12
US-09-647-344A-14
; Sequence 14, Application US/0964734A
; Patent No. 6586180
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong
; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678.PCT.US
; CURRENT APPLICATION NUMBER: US/09/647,344A
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: PCT/US99/06742
; PRIOR FILING DATE: 1999-03-28
; NUMBER OF SEQ ID NOS: 50
; SEQ ID NO 14
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pShuttle
US-09-647-344A-14

Query Match   6.5%;   Score 42.8;   DB 4;   Length 8705;
Best Local Similarity 58.7%;   Pred. No. 0.074;
Matches 74;   Conservative 0;   Mismatches 52;   Indels 0;   Gaps 0;

Qy   13 CTGCACCTTGGCCTCTGCAGCCACCTGCTGCCCCGGCCCCCTCTGGCCCCTTGAG 72
         ||||||||||| ||| |||||||||||| |||||||||||||||| ||||||||
Db 7305 CTCCACCTCCTGTCCTCCTGCCCCTCCTGCGGCCCCTCCTGCCCCTCCTG 7364

Qy   73 GCCCCCCAGCTTCTGCCTGTGCTGTGATCCCCCAGCCAGCGGGGCACGGCCCCGGCCTTCCGC 132
         |||||| ||| ||||||||| || |||||||||| ||||||||||||||||| |||||||
Db 7365 CCCTTCCTGCTCCTGCCCCTCCTGCCCCCTCGTCCCCTGCTCCCCTTCCTC 7424

Qy   133 TGCCCC 138
         ||||||
Db 7425 CTGCCC 7430

RESULT 13
US-08-910-647-1/c
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville

; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,647
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
US-08-910-647-1
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

Query Match   6.5%;   Score 42.8;   DB 3;   Length 9600;
Best Local Similarity 58.7%;   Pred. No. 0.077;
Matches 74;   Conservative 0;   Mismatches 52;   Indels 0;   Gaps 0;

Qy   13 CTGCACCTTGGCCTCTGCAGCCACCTGCTGCCCCGGCCCCCTCTGGCCCCTTGAG 72
         ||||||||||| ||| |||||||||||| |||||||||||||||| ||||||||
Db 1411 CTCCACCTCCTGTCCTCCTGCCCCTCCTGCGGCCCCTCCTGCCCCTCCTG 1352

Qy   73 GCCCCCCAGCTTCTGCCTGTGCTGTGATCCCCCAGCCAGCGGGGCACGGCCCCGGCCTTCCGC 132
         |||||| ||| ||||||||| || |||||||||| ||||||||||||||||| |||||||
Db 1351 CCCTTCCTGCTCCTGCCCCTCCTGCCCCCTCGTCCCCTGCTCCCCTTCCTC 1292

Qy   133 TGCCCC 138
         ||||||
Db 1291 CTGCCC 1286

RESULT 14
US-09-620-925-1/c
; Sequence 1, Application US/09620925
; Patent No. 6468986
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/620,925
; FILING DATE: 21-Jul-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/910,647
; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:
;   NAME: Fujita, Sharon M.
;   REGISTRATION NUMBER: 38,459
;   REFERENCE/DOCKET NUMBER: 1218.002
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (510) 923-2706
;   TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 9600 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1

Query Match   6.5%; Score 42.8; DB 4; Length 9600;
Best Local Similarity 58.7%; Pred. No. 0.077;
Matches   74; Conservative   0; Mismatches 52; Indels   0; Gaps   0;

Qy   13 CTGCACCTTCGCCTCTGCAGCCACCTTCGGCTGCCCCGGGCCCCCTTCGGCCCCCTTGAG 72
         ||||||||| ||||||||||||||||||||||||||||||||||||| |||| ||||||||
Db 1411 CTCCACCTTCCTCTGCTCCTGCCCCCTTCCGCCCCCTTCCGCTCCTGCCCCCTTCCTG 1352

Qy   73 GCCCCCCAGCTTCTGCCTGTGCTGTGATCCCCCACCCCCAGCCGCACGGGCCCCGCCTCCGC 132
         ||| ||||||| ||||||||| |||||||||||||||||||||||||||||||||||||||
Db 1351 CCCCTCCTGCTCCTCCGCCCCCTTCTGCCCCTTCCTGCCCCTTCCTGCCCCCTTCCTC 1292

Qy   133 TGCCCC 138
         ||| |||
Db 1291 CTGCCC 1286

RESULT 15
; Sequence 15, Application US/07884811
; Patent No. 5316921
; GENERAL INFORMATION:
;   APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
;   TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
;   NUMBER OF SEQUENCES: 21
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/884,811
;   FILING DATE: 19920518
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Dreger, Ginger R.
;   REGISTRATION NUMBER: 33,055
;   REFERENCE/DOCKET NUMBER: 755.1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-3216
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 10596 bases
;   TYPE: NUCLEIC ACID

;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-07-884-811-15

Query Match   6.5%; Score 42.8; DB 1; Length 10596;
Best Local Similarity 58.7%; Pred. No. 0.08;
Matches   74; Conservative   0; Mismatches 52; Indels   0; Gaps   0;

Qy   13 CTGCACCTTCGCCTCTGCAGCCACCTTCGGCTGCCCCGGGCCCCCTTCGGCCCCCTTGAG 72
         ||||||||| ||||||||||||||||||||||||||||||||||||| |||| ||||||||
Db 2945 CTCCACCTTCCTCTGCTCCTGCCCCCTTCCGCCCCCTTCCGCTCCTGCCCCCTTCCTG 2886

Qy   73 GCCCCCCAGCTTCTGCCTGTGCTGTGATCCCCCACCCCCAGCCGCACGGGCCCCGCCTCCGC 132
         ||| ||||||| ||||||||| |||||||||||||||||||||||||||||||||||||||
Db 2885 CCCCTCCTGCTCCTCCGCCCCCTTCTGCCCCTTCCTGCCCCTTCCTGCCCCCTTCCTC 2826

Qy   133 TGCCCC 138
         ||| |||
Db 2825 CTGCCC 2820

Search completed: May 22, 2004, 17:00:50
Job time : 83 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 22, 2004, 14:52:52 ; Search time 2049 seconds
                 (without alignments)
                 13876.540 Million cell updates/sec

Title: US-10-060-066-2
Perfect score: 656
Sequence: 1 tatggcttcactcgcacct..........tgtcatttcattctttg 656

Scoring table: IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :
                 GenEmbl:*
                 1: gb_ba:*
                 2: gb_htg:*
                 3: gb_in:*
                 4: gb_om:*
                 5: gb_ov:*
                 6: gb_pat:*
                 7: gb_ph:*
                 8: gb_pl:*
                 9: gb_pr:*
                 10: gb_ro:*
                 11: gb_sts:*
                 12: gb_sy:*
                 13: gb_un:*
                 14: gb_vi:*
                 15: gb_vl:*
                 16: gb_ba:*
                 17: gb_fun:*
                 18: em_in:*
                 19: em_mu:*
                 20: em_om:*
                 21: em_or:*
                 22: em_ov:*
                 23: em_pat:*
                 24: em_ph:*
                 25: em_pl:*
                 26: em_ro:*
                 27: em_sts:*
                 28: em_un:*
                 29: em_vi:*
                 30: em_vrt:*
                 31: em_htg_inv:*
                 32: em_htg_other:*
                 33: em_htg_mus:*
                 34: em_htg_pln:*
                 35: em_htg_rod:*
                 36: em_htg_mam:*
                 37: em_htg_vrt:*
                 38: em_sy:*
                 39: em_htgo_hum:*
                 40: em_htgo_mus:*
                 41: em_htgo_other:*

† Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	592.8	90.4	185839	2	AC079557	AC079557 Mus muscu
2	519.4	79.2	244548	2	AC128446	AC128446 Rattus no
3	519.4	79.2	246624	2	AC114434	AC114434 Rattus no
4	387.4	59.1	1380	10	AF361350	AF361350 Mus muscu
5	297.8	45.4	169997	9	AC004440	AC004440 Homo sapi
6	269.8	41.1	1266	10	AF361346	AF361346 Rattus no
7	236	36.0	190210	2	AC022318	AC022318 Homo sapi
8	232.2	35.4	1386	9	AF361354	AF361354 Homo sapi
9	191.2	29.1	633	9	AF234892	AF234892 Homo sapi
10	171	26.1	1281	9	AF288388	AF288388 Homo sapi
11	92	14.0	987	5	AY037891	AY037891 Gallus ga
12	86.4	13.2	720	9	AF142622	AF142622 Homo sapi
13	86.4	13.2	984	6	AR165149	AR165149 Sequence
14	86.4	13.2	984	10	AX101260	AX101260 Sequence
15	86.4	13.2	984	10	MMU272045	AJ272045 Mus muscu
16	86.4	13.2	1001	9	AF361341	AF361341 Rattus no
17	86.4	13.2	3392	6	AF162692	AF162692 Homo sapi
18	86.4	13.2	94538	10	BC034532	BC034532 Homo sapi
19	86.4	13.2	173126	9	AL645853	AL645853 Mouse DNA
20	86.4	13.2	178380	9	AC005988	AC005988 Homo sapi
21	86.4	13.2	1071	6	AC129819	AC129819 Rattus no
22	83.2	12.7	209887	6	AR202726	AR202726 Sequence
23	75.4	11.5	209897	2	AC079424	AC079424 Human DNA
24	61.4	9.4	1548	10	AF118819	AF118819 Rattus no
25	61.4	9.4	3067	5	BC057439	BC057439 Danio rer
26	58.6	8.9	125020	9	AF429315	AF429315 Homo sapi
27	57.6	8.8	196036	9	BX004779	BX004779 Zebrafish
28	56.2	8.6	203225	5	BX255942	BX255942 Zebrafish
29	54.6	8.3	1266667	2	AC015839	AC015839 Homo sapi
30	54.4	8.3	972	2	AF096322	AF096322 Homo sapi
31	54.4	8.3	70881	9	BX284660	BX284660 Human DNA
32	54.4	8.3	99497	9	HS291L6	HS291L6 Human DNA
33	53.6	8.2	948	10	AL047749	AL047749 Human DNA
34	53.6	8.2	948	10	MMU272044	AJ272044 Mus muscu
35	53.6	8.2	948	10	AF361340	AF361340 Rattus no
36	53.6	8.2	191935	10	AR202725	AR202725 Sequence
37	53.6	8.2	237373	10	AC016522	AC016522 Mus muscu
38	53.6	8.1	252720	2	AC094178	AC094178 Rattus no
39	53.2	8.1	63082	2	AC132248	AC132248 Mus muscu
40	52.8	8.0	972	10	AC022663	AC022663 Homo sapi
41	52.8	8.0	1124	10	AF361339	AF361339 Rattus no
42	52.8	8.0	1554	10	AF118818	AF118818 Rattus no
43	52.8	8.0	1558	6	AF077739	AF077739 Mus muscu
44	52.8	8.0	197660	10	AR202724	AR202724 Sequence
45	52.8	8.0	208651	10	AL589650	AL589650 Mouse DNA
					AC090493	AC090493 Genomic 8

ALIGNMENTS

RESULT 1
AC079557/c
LOCUS   AC079557   185839 bp   DNA   linear   HTG 02-SEP-2000
DEFINITION   Mus musculus clone RP23-449F12, WORKING DRAFT SEQUENCE, 16
              unordered pieces.
ACCESSION   AC079557
VERSION   AC079557.1 GI:9964922
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 185839)
AUTHORS   DOE Joint Genome Institute.
TITLE   Sequencing of Mouse

JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 185839)
  AUTHORS   DOE Joint Genome Institute.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
             Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT   -------------Genome Center
             Center: Joint Genome Institute
             Center Code: JGI
             Web site: http://www.jgi.doe.gov

             Project Information
             Center Project Name: 1909420
             Center clone name: RPCI-23_449F12

             Summary Statistics
             Consensus quality: 172503 bases at least Q40
             Consensus quality: 177768 bases at least Q30
             Consensus quality: 179073 bases at least Q20
             Estimated insert size: 187000; agarose-fp estimation
             Estimated insert size: 184339; sum-of-contigs estimation
             Quality coverage: 7.81 in Q20 bases; agarose-fp estimation
             Quality coverage: 7.93 in Q20 bases; sum-of-contigs estimation.
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 16 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *
             *   1   1081: contig of 1081 bp in length
             *   1082   1181: gap of unknown length
             *   1182   4298: contig of 3117 bp in length
             *   4299   4398: gap of unknown length
             *   4399   6732: contig of 2334 bp in length
             *   6733   6832: gap of unknown length
             *   6833   11626: contig of 4794 bp in length
             * 11627   11726: gap of unknown length
             * 11727   15545: contig of 3819 bp in length
             * 15546   15645: gap of unknown length
             * 15646   20007: contig of 4362 bp in length
             * 20008   20107: gap of unknown length
             * 20108   26156: contig of 6049 bp in length
             * 26157   26256: gap of unknown length
             * 26257   32958: contig of 6702 bp in length
             * 32959   33058: gap of unknown length
             * 33059   42014: contig of 8956 bp in length
             * 42015   42114: gap of unknown length
             * 42115   50688: contig of 8574 bp in length
             * 50689   50788: gap of unknown length
             * 50789   61298: contig of 10510 bp in length
             * 61299   61398: gap of unknown length
             * 61399   77269: contig of 15871 bp in length
             * 77270   77369: gap of unknown length
             * 77370   95693: contig of 18324 bp in length
             * 95694   95793: gap of unknown length
             * 95794   119999: contig of 24206 bp in length
             * 120000   120099: gap of unknown length
             * 120100   150273: contig of 30174 bp in length
             * 150274   150373: gap of unknown length
             * 150374   185839: contig of 35466 bp in length.
FEATURES   Location/Qualifiers
     source   1..185839
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-449F12"
                     /clone_lib="RPCI mouse BAC library 23"

ORIGIN
Query Match   90.4%; Score 592.8; DB 2; Length 185839;
Best Local Similarity 95.3%; Pred. No. 2e-129;

Matches 625; Conservative 0; Mismatches 29; Indels 2; Gaps 2;

```
Qy       1 TATGGCTTCACTCTGCACCTTCGCCTTCGCCTGCAGCCACCTGCTCGGCTCGCCCCGGCCCCCCTC 60
           |||||||||||||||||||||||| |||||||||||| |||||||||||||||||||||||||||||
Db   57586 TATGGCTTCACTCTG-ACCTTCGCCTTCGCCTTCTGCAGCCACCTGCTCGGCTCGCCCCGGCCCTC 57528

Qy      61 TGGCCCCTTGAGGCCCCAGCGTTCTGCGTGCTGTCTGATTCCCCCACCGCGCCCCACCGCACCGG 120
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   57527 TGGCCCCTTGAGGCCCCAGCTTCTGCGTGCTGTCTGATTCCCCCACCGCGCCCCACCGCACCGG 57468

Qy     121 CCCGGCCTTCCGCGTGGCCCCGGGGCCCCACCAGCTTCGCCTGTGGCTCGGCGATCCCCCACCGG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57467 CCCGGCCTTCCGCGTGGCCCCGGGGCCCCACCAGCTTCGCCTGTGGCTCGGCGATCCCCCACCGG 57408

Qy     181 AGTCATTGAAACCGCTGGAATGAAGAGAGGGGTTTGGTGGTGTGGTGTGAAAAGGGCGTTCAGGTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57407 AGTCATTGAAACCGCTGGAATGAAGAGAGGGGTTTGGTGGTGTGGTGTGAAAAGGGCGTTCAGGTAC 57348

Qy     241 TACTGACACCATCGGCCGCCCCCTTCGCGGCGTTTTTGCGGCCCTCATGACCATCCCCATCAGCACTG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57347 TACTGACACCATCGGCCGCCCCCTTCGCGGCGTTTTGGCCCTCATGACCATCCCCATCAGCACTG 57288

Qy     301 ACTACTCGGCTCTTACACAAGAGAGCTCTCATCTGCAACACACCACCTCACAGCAGGTGATG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57287 ACTACTCGGCTCTTACACAAGAGACTTCTCATCTGCAACACACCACCTCACAGCAGGTGATG 57228

Qy     361 ACGGACCACCCCATCGTGGGGGGCCAGTGGCTCCTCCGAGAAGAAGGACCCCGGGGGGCCTCA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57227 ACGGACCACCCCATCGTGGGGGGCCAGTGGCTCCTCCGAGAAGAAGGACCCCTGGGGGGCCTCA 57168

Qy     421 CACATTCAGGCCTTCTGCGGGGGATATGCTGCCCTGGAAGGTGAGGGGTGCAGAAGGCCCTGGCT 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57167 CACATTCAGGCCTTCTGCGGGGGATATGCTGCCCTGGAAGGTGAGGGGTGCAGGAAGGCCCTGGCT 57108

Qy     481 GGCCACCTCCGNTGAAGCCTGTGCCTTNCAGCAAAATCGACTTNTAGTTGTTNCTGGTCTCAA 540
           ||||||||| ||| |||||||||||||||||||| |||||||||||||||||||||||||||||
Db   57107 GGCCACCTCCGCTGAAGCCTGTGCCTCCAGCAAAATCGACTTNTAGTTGTTCTGGTCTCAT 57048

Qy     541 CCCAAGGNCTTGAGGGGTGTGCTTTACTGCTGGAGGCGGNNRATTCCATGCACCCCTCCA 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57047 CCCAAGGTCTTGAGGGGTGTGCTTTACTGCTGGAGGCGGGGATTCCATGCACCCCTCCA 56988

Qy     601 TNCCTNTNTACAAAATCCCTCTTCTCTGGGGATTNTNGNGNTGTTTATTTTTATTTTCTTTG 656
           ||||| |||||||||||||||||||||||||||||||||| |||||| |||||||||||||||
Db   56987 TCCCTCTGT-CCAATCCCTCTTCTCTGGGGATTCTCGTTTATTTAATTTTATACTGG 56933
```

RESULT 2
LOCUS   AC128446/c
ACCESSION   AC128446
LOCUS   AC128446   244548 bp   DNA   linear   HTG 09-OCT-2002
DEFINITION Rattus norvegicus clone CH230-95B19, *** SEQUENCING IN PROGRESS
            ***.
ACCESSION   AC128446
VERSION   AC128446.2 GI:22856047
KEYWORDS   HTG; HTGS_PHASE2; HTGS DRAFT; HTGS_ENRICHED.
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 244548)
  AUTHORS   Muzny,D.Marie., Metzker M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aovagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Dunn,K., Duval,B., Eaves,K.,

```
                                    Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
                                    *  of the sequence length compared to sum-of-contigs length
                                    *  NOTE: Estimated insert size may differ from sequence length
                                    *  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
                                    *  NOTE: This is a 'working draft' sequence. It currently
                                    *  consists of 1 contigs. Gaps between the contigs
                                    *  are represented as runs of N. The order of the pieces
                                    *  is believed to be correct as given, however the sizes
                                    *  of the gaps between them are based on estimates that have
                                    *  provided by the submitter.
                                    *  This sequence will be replaced
                                    *  by the finished sequence as soon as it is available and
                                    *  the accession number will be preserved.
                                    *  1  244548; contig of 244548 bp in length.
FEATURES             Location/Qualifiers
     source          1..244548
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-95B19"
ORIGIN
Query Match  79.2%;  Score 519.4;  DB 2;  Length 244548;
Best Local Similarity 94.0%;  Pred. No. 4.4e-112;
Matches 568;  Conservative  0;  Mismatches 33;  Indels 3;  Gaps 3;
```

```
Qy       2 ATGGCTTCACTCTGCACCTTCGCCTCTCGAGCCACCTGCTCGCTGCCCCCGGCCCCTCT      61
           ||||||||||||| ||||| ||||||| |||||||||||||||||||| ||||||||||||
Db  100974 ATGACTTCACTCTG-ACCTCCGCCTCTCTGCAGCCACCTGCTCGCTGCCCCCTT-CCCCTCT 100917

Qy      62 GGCCCCTTGAGGCCCCCCAGCTTCTTGCCTGTGCGTGATCCCCCCCAGCCGCCCGCACGGC     121
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  100916 GGCCCCTCGAGGCCCCCCAGCTTCTTGCCTGTGCGTGATCCCCCCCAGCCGCCCGGCACGGC 100857

Qy     122 CCCCGCCTCCCCGTGGGTGGGTGGCCCCCACGGCCCCCCGGTTGCCCAGTGGGTGAAACTGGA     181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  100856 CCCCGCCTCCCCGTGGGTGGGTGGCCCCCACGGCCCCCCGGTTGCCCAGTGGGTCAAACTGGA 100797

Qy     182 GTCATTGAAACGCTGGAATGAAGAGAGAGGGGGTTTGTGGGTGGTGAAAAGGGGCGGTTACT     241
           ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  100796 ATCATTGAAACGCTGGAATGAAGAGAGAGGGGGTTTGTGGGTGGTGAAAAGGGGCGGTTACT 100737

Qy     242 ACTGACACCATCGGCGCGCCTTCGCGGGCCTTTGGGCCTCCATGACCAATCGACCATCAGCACTGA     301
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  100736 ACTGACACCATAGGCGCCCTTTGCAGCTTTGGGCCTCATGACCAATCGACCATCAGCACTGA 100677

Qy     302 CTACTGGGCTCTACACAAGAGACTTCTCATCTGCAACACCACCCAACCTCACACAGCAGGTGATGA     361
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  100676 CTACTGGGCTCTACACAAGAGACTTCTCATCTGCAACACCACCCAACCTCACACAGCAGGTGATGA 100617

Qy     362 CGGACACCCCATCGTGGGCGGCAGTGGCCTCCTCCGAGAGAAGAGAGGACCCTGGGGGGCCTCAC     421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  100616 TGGACCACCCCATCGTGGGCGGCAGTGGCCTCCTCAGAGAGAAGAGAGGACCCTGGGGGGCCTCAC 100557

Qy     422 ACAATTCAGGCGCCTCTGGGCCGGATATGCTGCCTGGAGGTAGGGTGCAGGAGGACCCCTGGGCTG     481
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  100556 ACATTCAGGCGCCTCTGCGGCGGATATGCTGCCTGGACAGGTAGGGTGCAGGAGGACCCAGGGCTG 100497

Qy     482 GCCACCTCCGNTGAAGCCTGTGCCTCNCAGCAAAATGCACTTNTAGTTGTTCTGAAGTCTCAAC     541
           |||||||||||||||||||||||||| ||||| ||||||||||| |||||||||||||||||||
Db  100496 GCCACCTCCGCGTGAAGCCTGTGCCTCTCAGCAAAAGCGACTTGCTAGTTGTTCTGAAGTCTCATC 100437

Qy     542 CCAAGGNCTTGAGGGGTGTCCTTTACTGCTGGAGGCG-GNTNATTCCATGCACCCCTCCA     600
           ||||| ||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  100436 CCAAGGTCTTGAGGGGGTGTGCTTCACTGCTGGAGGAGCTATGATTCCATGCACCCCTCCTCTG 100377

Qy     601 TNCC     604
           ||||
Db  100376 TCCC 100373
```

                    Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
                    Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
                    Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
                    Gebregeorgis,E., Geer,R., Gill,P., Grady,M., Guerra,W.,
                    Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
                    Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
                    Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
                    Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
                    Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
                    Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
                    Kowis,C., Kraft,C.L., Lebow,H., Leven,J., Lewis,L., Li,Z., Liu,J.,
                    Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
                    Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
                    Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
                    Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
                    Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
                    Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
                    Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
                    Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
                    Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pai,S., Parks,K.,
                    Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
                    Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
                    Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
                    Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
                    Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
                    Sanders,W., Savery,G., Scherer,S., Scott,G., Shateman,S., Shen,H.,
                    Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Small,D.,
                    Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
                    Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
                    Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
                    Valdes,V., Vera,V., Villasana,D., Waldron,J., Walker,B., Wang,J.,
                    Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
                    Williams,G., Willson,R., Wieczyk,R., Wooden,H., Worley,K.,
                    Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,V.,
                    Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
                    Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
                    Weinstock,G. and Gibbs,R.A.
  JOURNAL   Direct Submission
            Unpublished
REFERENCE   2 (bases 1 to 244548)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 244548)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE   Direct Submission
  JOURNAL   Submitted (09-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Sep 14, 2002 this sequence version replaced gi:21901138.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads. Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            -------------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            -------------------- Project Information
            Center project name: GYWS
            Center clone name: CH230-95B19
            -------------------- Summary Statistics
            Assembly program: Phrap; Version 0.990329
            Consensus quality: 206685 bases at least Q40
            Consensus quality: 209611 bases at least Q30
            Consensus quality: 211454 bases at least Q20
            Estimated insert size: 227005; sum-of-contigs estimation

RESULT 3
AC114434/c

```
LOCUS       AC114434              246624 bp    DNA     linear   HTG 21-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-210C18, *** SEQUENCING IN PROGRESS
            ***.
ACCESSION   AC114434
VERSION     AC114434.3  GI:23266184
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 246624)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Gantt,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,C., Guevara,W.,
            Gunaratne,P., Haaland,M., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,B., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindaraine,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarunpunsagoon,A., Pai,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,S.J.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steinle,M., Strong,R., Sutton,G., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Vales,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Wilson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 246624)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 246624)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 21, 2002 this sequence version replaced gi:21737492.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads.  Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            ---------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
            Center project name: GPFX
            Center clone name: CH230-210C18
            ---------- Summary Statistics
            Assembly program: Phrap; version 0.990329
            Consensus quality: 214395 bases at least Q40
            Consensus quality: 216765 bases at least Q30
            Consensus quality: 218383 bases at least Q20
            Estimated insert size: 238927; sum-of-contigs estimation
            Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
            ------------------------------------------------------
            * NOTE: Estimated insert size may differ from sequence length
            * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            * 1  246624: contig of 246624 bp in length.
FEATURES             Location/Qualifiers
     source          1..246624
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-210C18"
     misc_feature    144..702
                     /note="clone_boundary"
                     /site:EcoRI
                     clone_end:T7
                     end_sequence:BH344045"
     misc_feature    complement(244274..244949)
                     /note="clone_boundary"
                     /site:EcoRI
                     clone_end:Sp6
                     end_sequence:BH344045"
ORIGIN
```

```
Query Match
Best Local Similarity   94.0%;
Matches   568;  Conservative   0;  Mismatches   33;  Indels   3;  Gaps   3;

Qy         2 ATGGCTTCACTCTGCACCTTCGCCTCTGCAGCCACCTGCTCGCTTGCCCCGGCGCCCCTCT   61
             ||| |||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||
Db    174382 ATGACTTCACTCTG-ACCTCCGCCTCTGCAGCCACACTGCTCGCTTGCCCCCTCT-CCCCTCT 174325

Qy        62 GGCCCCTTGAGGCGCCCCCAGCTTCTGCCTGTGTCTGTGATCCCCCAGCGCGCCCGGCACGGC 121
             |||||| ||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    174324 GGCCCCTCCGAGGCGCCCCCAGCTTCTGCCTGTGTCTGTGACCCCCCAGCGCGCCCGGCACGC 174265

Qy       122 CCCGGCCTCGCTGCCCCGGGTGGTGGCCACGGCGCCCCCGGTTGCCAGTGGTGAAACTGGA 181
             |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    174264 CCCGGCCTCGCTGCCCCGGGTGGTGGCCACGGCGCCCCCGGTTGCCCGTGGTCAAACTGGA 174205

Qy       182 GTCATTGAAACCCTGGAATGAGAGAGGGGGGTTTGTGGTGTGAAAAGGGCGTTCAGGTACT 241
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db 174204 ATCATTGAAACGCTGGGAATGAAGAGAGAGGGGTTTGTGGTGCGGAAAAGGGCGTTCAGGTACT
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 242 ACTGACACCACCATCGGCGCGCCTTCGCGGGGCTTTTGGGCCTCATGACCATCAGCACTGA 301
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 174144 ACTGACACCACCATAGGCGCGCCTTTGCAGCTTTTTGGCCTCATCGACCATCAGCACTGA 174085

Qy 302 CTACTGGGTCTCTACAACAGAGCTCTCATCTGCACAACAGCAGCAGGTGATGA 361
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 174084 CTACTGGGTCTCTACAACAGAGCTCTCATCTGCACAACACCACCAACCTCACAGCAGCAGGTGATGA 174025

Qy 362 CGGACCAACCCCATCGTGGGGGGGCAGTGGTCCTCTCCGAGAAGGGGAAGGACCCCTGGGGGGCCTCAC 421
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 174024 TGGACCAACCCCATCGTGGGGGGGCAGTGGTCCTCTCCGAGAAGGGGAAGGACCCCTGGGGGGCCTCAC 173965

Qy 422 ACATTCAGGCGCCTCTGGCGGATATCGTGCTGGGAAGGTAGGGTGCAGGAGGGCCCTGGCTG 481
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 173964 ACATTCAGGCGCCTCTGGCGGATATCGTGCTGGGAAGGTAGGGTGCAGGAGGGCCCTGGCTG 173905

Qy 482 GCCACCTCCGNTGAAGCCTGTGCCTNCAGCAAAATCGACTTNAGTGTTCTCAAC 541
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 173904 GCCACCTCCGCTGAAGCCTGTGCCTCCAGCAAATCGACTGCTAGTGTTCTCAAC 173845

Qy 542 CCAAGGNCTTGAGGGGTGTGTCTTTACTGCCTCCACGAAAACGACTAGTGTTCTGAANTCTCATC 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 173844 CCAAGGTCTTGAGGGGTGTGTCTTCACTGCTGTGGAGGAGCTAGTGCATCCTCCTCTG 173785

Qy 601 TNCC 604
        ||||
Db 173784 TCCC 173781

RESULT 4

LOCUS   AF361350   1380 bp   mRNA   linear   ROD 21-DEC-2001
DEFINITION Mus musculus voltage-dependent calcium channel gamma-8 subunit
            (Cacng8) mRNA, complete cds.
ACCESSION   AF361350
VERSION   AF361350.1 GI:17974533
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1380)
  AUTHORS   Chu,P.J., Robertson,H.M. and Best,P.M.
  TITLE   Calcium channel gamma subunits provide insights into the evolution
            of this gene family
  JOURNAL   Gene 280 (1-2), 37-48 (2001)
   MEDLINE   21601102
   PUBMED   11738816
REFERENCE   2 (bases 1 to 1380)
  AUTHORS   Chu,P.-J., Robertson,H. and Best,P.M.
  TITLE   Direct Submission
  JOURNAL   Submitted (14-MAR-2001) Department of Molecular and Integrative
            Physiology, University of Illinois, 407 S. Goodwin Ave., Urbana, IL
            61801, USA
FEATURES   Location/Qualifiers
     source   1..1380
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
     gene   1..1380
                     /gene="Cacng8"
     CDS   109..1380
                     /gene="Cacng8"
                     /codon_start=1
                     /product="voltage-dependent calcium channel gamma-8
                     subunit"
                     /protein_id="AAL50045.1"
                     /db_xref="GI:17974534"
                     /translation="MESLKRWNEERGLWCEKGVQVLLTTIGAFSAFGLMTIAISTDYW"

ORIGIN
Query Match   59.1%; Score 387.4; DB 10; Length 1380;
Best Local Similarity   99.7%; Pred: No. 6e-81;
Matches   388; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   69 TGAGGGCCCCCCAGCTTCTGCGTGCTGTGATCCCCACCCCCAGCCGCACGGGCCCGCGCCT 128
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGAGGGCCCCCCAGCTTCTGCGTGCTGTGATCCCCACCCCCAGCCGCACGGGCCCGCGCCT 60

Qy 129 CCGCTGCGCCCCGGTGGTGGCCCACGGGCCCCCCCCCGGTTGGTCGAAACTGGAGTGCATTG 188
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCGCTGCGCCCCGGTGGTGGCCCACGGGCCCCCCCCCGGTTGGTCGAAACTGGAGTGCATTG 120

Qy 189 AAAACCCTGGAATGAAGAGAGAGGGGTTTGTGGTGTGTGAAAAGGGCGTTCAGGTACTACTGACC 248
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AAAACCGCTGGAATGAAGAGAGAGGGGTTTGTGGTGTGTGAAAAGGGCGTTCAGGTACTACTGACC 180

Qy 249 ACCATCGGCGCCTTCGGCGTTTGGCCTTTNGGCCTCATGACCATCGGCCATCGACTGACTACTGG 308
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 ACCATCGGCGCCTTCTCGGCTTTCGGCCTTTTGGCCTCATGACCATCAGCACTGACTACTGG 240

Qy 309 CTCTACACAGAGACTCTCATCTGCAACACAACCACCAACCTCACAGCAGCAGGTGATGACGGACCA 368
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 CTCTACACAGAGACTCTCATCTGCAACACAACCACCAACCTCACAGCAGCAGGTGATGACGGACCA 300

Qy 369 CCCCATCGTGGGGGGCGCAGTGGTGCTCTTCTCCGAGAAGAAGGAACCCCTGGGGGCCTCACAATTCA 428
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CCCCATCGTGGGGGGCGCAGTGGTGCTCTTCTCCGAGAAGAAGGAACCCCTGGGGGCCTCACAATTCA 360

Qy 429 GGCCTCTGGCGGATATGCTGCCCTGGAAGG 457
        ||||||||||||||||||||||||||||||
Db 361 GGCCTCTGGCGGATATGCTGCCCTGGAAGG 389

RESULT 5

LOCUS   AC008440   169997 bp   DNA   linear   PRI 13-JUL-2002
DEFINITION Homo sapiens chromosome 19 clone CTC-331H23, complete sequence.
ACCESSION   AC008440
VERSION   AC008440.9 GI:21743728
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 169997)
  AUTHORS   DOE Joint Genome Institute.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 169997)
  AUTHORS   DOE Joint Genome Institute.
  TITLE   Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3 (bases 1 to 169997)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE   Direct Submission
  JOURNAL   Submitted (01-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4 (bases 1 to 169997)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-JUL-2002) DOE Joint Genome Institute and Stanford Human Genome Center.
            DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT   On Jul 13, 2002 this sequence version replaced gi:14277251.

LYTRALICNTTNLTAGDDGPHRGGSGSSSEKDPGGLITHSGLWRICLLEGLKRGVCVK
INHFPEDIYDYDHDSAEYLLRVVRASSIFPILSAILLLGGVCVAASRVYKSKRNILG
AGILPVAAGLSNIIGVIVISANAGEPGPKRDEBEKKNHYSYGNSFYFGGLSFILAEVI
GVLAVNIYIERSREAHCQSRSDLLKAGGGAGGSGGSGSPSAILRLPSYRPYRYRSRSS
SRGSSEASPSRDASPGGEPGGPGFRASTDISMYTLSRDPSKGSVAAGLASAGGSGSGAGV
GAYGGAGAAGGGGAGSERDRGSSAGFLTIHNAPPKEAASGVTVTVTGFPAAPAPAPA
PPAPAAPAPGTLSKEAAASNTNTLNRKTTPV"

```
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.3.

NOTE: 90954 Ambiguous base. Probably T.
WI-17997 G23480
STS Content:
```

FEATURES
    source
```
Location/Qualifiers
1. .169997
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="19"
    /clone="CTC-331H23"
```

ORIGIN

Best Local Match

Query Match
Best Local Similarity   81.2%;   Score 297.8;   DB 9;   Length 169997;
Matches 384;   Conservative   0;   Mismatches 82;   Indels 7;   Gaps 3;

```
Qy       9 CACTCTGCACCTTCGCCTTCTGCAGCCACCTGCTGCTGCCCCGCCCCC-CTCTCGGCGCCCC 67
             |||||||||||||||||||||||||||||||||| |||||||||||||||  ||||||||||||
Db  107335 CTCTGACCCCTTCCCCGGCGCCGAGCCTTCCTCTCTCGCCGCCGCCGCCGCCCCGCCG 107394

Qy      68 TTGAGGCGCCCCCCCAGCTTCTTCTGCTTGTGCTTGTGAT---CCCCCAGCCGCCGCCGCCGCCCC 124
             |||||||||||||||||||| |||||||||||| |||||   |||||||||||||||||||||||
Db  107395 CCCCGGCGCCCCCGGCTTCTGCCTGCGGTGAACCCCCCCCAGCCCGCCGCACGGGCCCC 107454

Qy     125 GCCTCCGCCTGCCCCGGTGGTGGGCCCACGGGCCCACGGGCCCCCCCCGGTGGTGGTGAAAACTGGAGTC 184
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107455 GCCCCCGGCCTGCCCCGGTGGTGGGCCCACGGGCCCCCCCCGGTGGTGGTGCAGGTACTACTACT 107514

Qy     185 ATTGAAAGGCTGGAAATGAAGAGAGAGGGGGTTTGTGGTTGTGAAAAGGCGGTTCAGGTACTACTACT 244
             ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107515 GCTGAAGCGCTGGAAACAGAAGAGCGGGGGCCCTTCTGGGTGGGCAGGGGGTGCAGGTGTGCTGCT 107574

Qy     245 GACCACCATCGCGCGCTTCGCGGCCTTTTGGGCTCCTCATCTGCCCATCGCCATCAGCACTCGACTA 304
             ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  107575 GACGACCGGTGGCGCCTTCGCCGCCCCTTGGGCCTGGGCCTTGGGGGTTGTGCGCGAACACTGACTA 107634

Qy     305 CTGGCCTTACACAAGAGCTCTCATCTGCAACACCACCAACCTCACAGC---AGGTGATGA 361
             ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db  107635 CTGGCTCTACACGGCGGCGCCCCTCCAACACCACCAACCTTCACGGCGTTGGGGGGCGCGGA 107694

Qy     362 CGGACCACCCATCGTGGGGGCAGTGGCTCTCCGAGAAGAAGAGGACCCCTGGGGCGGCCTCAC 421
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107695 CGGGACCCCCCACGGGGGCGGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGCGCTCAC 107754

Qy     422 ACATTCAGGCCTTCTGGGGATATGCTCGCCTGGGAAGGTAGGGTGGTTGCAGGAAGGCC 474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107755 GCACTCGGGCCCTTCTGGGAGGATCTGCTCGCCTGGGAAGGTAGGGTCCGGCGGGGCGGCCC 107807
```

REFERENCE
AUTHORS
TITLE
JOURNAL

MEDLINE   21601102
PUBMED   11738816
REFERENCE 2 (bases 1 to 1266)
AUTHORS   Chu,P.-J., Robertson,H. and Best,P.M.
TITLE   Direct Submission
JOURNAL   Submitted (14-MAR-2001) Department of Molecular and Integrative
          Physiology, University of Illinois, 407 S. Goodwin Ave., Urbana, IL
          61801, USA

FEATURES
    source
```
Location/Qualifiers
1. .1266
    /organism="Rattus norvegicus"
    /mol_type="mRNA"
    /strain="Sprague-Dawley"
    /db_xref="taxon:10116"
```
    gene
```
1. .1266
    /gene="Cacng8"
```
    CDS
```
1. .1266
    /gene="Cacng8"
    /codon_start=1
    /product="voltage-dependent calcium channel gamma-8
    subunit"
    /protein_id="AAL50041.1"
    /db_xref="GI:17974526"
    /translation="MESLKRWNEBERGLWCEKGVQVLLTTIGAFAAFGLMTIAISTDYW
    LYTRALICNTTNLTAGDDGPHRGGSGSSEKKDPGGLTHSGLWRICLCELGRGVCVK
    INHFPEDTDYDHDSAEYLLRVVRAASSIFPILSAILLLLGGVCVAASRVYKSRNIILG
    AGILFVAAGLSNIIGVIVYISANAGEPGPKRDEEKKNHYSYGWSFYFGGLSFILAEVI
    GVLAVNIYIERSREAHCQSRSDLLKAGGAGGSGGSSGGSPSAILRLPSYRFYRFRRSRSS
    SRGSSEASPSRDASPGGPGGPGFASTDISMYTLSRDPSKGSVAAGLASAGGGGGGACV
    GAYGGAAGGAGGGGTGSSERDRGSSSAGFLTLHNAFPKEAASGVTVTVTGPPAAPAPAPP
    APAAPAPGTLSKEAAASNTNTLNRKTTPV"
```

ORIGIN

Best Local Match

Query Match
Best Local Similarity   97.5%;   Score 269.8;   DB 10;   Length 1266;
Matches 274;   Conservative   0;   Mismatches 7;   Indels 0;   Gaps 0;

```
Qy     177 CTGGAGTCATTGAAAACCGTGAAATGAAGAGAGAGGGGGTTTGTGGTTGTGAAAAGGCGGTTCAG 236
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTGGAGTCGTCATTGAAAACGGCTGGAAATGAAGAGAGAGGGGGTTGTGTGGTTGTGCGGAAAGGCCGGTTCAG 60

Qy     237 GTACTACTGACCACCATCGGGCGCGGCCCTTCGCGGCCTTTTGGGCCTCATGACCATCGCCATCAGC 296
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTACTACTGACCACCATAGGGCGCTTTTGCAGCTTTTTGCCCTCCATCATGACCATCGCCATCAGC 120

Qy     297 ACTGACTACTGGCTCTACACAAGAGCTCTCATCTGCAACACCACCAACCTCACAGCAGGT 356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ACTGACTACTGGCTCTACACAAGAGCTCTCATCTGCAACACCACCAACCTCACAGCAGGT 180

Qy     357 GATGACGGACCACCCCCATCGTGGGGGGCAGTGGGCTCCTCCGGAGAAGAAGAAGGAGGGGC 416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GATGATGGACCACCCCATCGTGGGGGGCAGTGGCTCCTCCGAGAAGAAGAAGGACCCTGGGGGC 240

Qy     417 CTCACACACATTCAGGCCTTCTGGGGGGATATGCTCGCCTGGGAAGG 457
             |||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTCACACATTCAGGCCTTCTGGGGGATATGCTCGCCTGGGAAGG 281
```

FEATURES
    source
```
Location/Qualifiers
1. .190210
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
```

RESULT 7
AC022318
LOCUS   AC022318   190210 bp   DNA   linear   HTG 31-AUG-2001
DEFINITION Homo sapiens chromosome 19 clone RP11-158G19, WORKING DRAFT
            SEQUENCE, 16 unordered pieces.
ACCESSION   AC022318
VERSION   AC022318.5 GI:15321555
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 190210)
AUTHORS   Waterston,R.H.
TITLE   The sequence of Homo sapiens clone
```

RESULT 6
LOCUS       AF361346   1266 bp   mRNA   linear   ROD 21-DEC-2001
DEFINITION  Rattus norvegicus voltage-dependent calcium channel gamma-8 subunit
            (Cacng8) mRNA, complete cds.
ACCESSION   AF361346
VERSION     AF361346.1  GI:17974525
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1266)
AUTHORS     Chu,P.-J., Robertson,H.M. and Best,P.M.
TITLE       Calcium channel gamma subunits provide insights into the evolution
            of this gene family
JOURNAL     Gene 280 (1-2), 37-48 (2001)
```

```
JOURNAL
REFERENCE   2  (bases 1 to 190210)
AUTHORS     Watereton,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (30-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Aug 28, 2001 this sequence version replaced gi:8516165.

            ------------- Genome Center -----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc/index.shtml
            ------- Project Information --------
            Center project name: H_NH0158G19

            ------------ Summary Statistics ------------
            Sequencing vector: M13; 64%
            Sequencing vector: plasmid; 36%
            Chemistry: Dye-primer ET; 64% of reads
            Chemistry: Dye-terminator Big Dye; 36% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 179489 bases at least Q40
            Consensus quality: 183023 bases at least Q30
            Consensus quality: 184976 bases at least Q20
            ------------------------------------
            NOTE: This is a 'working draft' sequence. It currently
            consists of 16 contigs. The true order of the pieces
            is not known and their order in this sequence record is
            arbitrary. Gaps between the contigs are represented as
            runs of N, but the exact sizes of the gaps are unknown.
            This record will be updated with the finished sequence
            as soon as it is available and the accession number will
            be preserved.
```

```
FEATURES             Location/Qualifiers
     source          1..190210
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="RP11-158G19"
     misc_feature    1..1029
                     /note="assembly_name:Contig14"
     misc_feature    1030..3107
                     /note="assembly_name:Contig17"
```

The following lines list contigs and gaps:

```
     1030:  contig  of 1029 bp in length
     1129:  gap     of unknown length
     1130:  contig  of 1978 bp in length
     3107:  gap     of unknown length
     3108:  contig  of 1548 bp in length
     3208:  gap     of unknown length
     4755:  contig  of 1548 bp in length
     4756:  gap     of unknown length
     4855:  contig  of 2232 bp in length
     4856:  gap     of unknown length
     7087:  contig  of unknown length
     7088:  gap     of unknown length
     7187:  contig  of 2575 bp in length
     7188:  gap     of unknown length
     9762:  contig  of unknown length
     9763:  gap     of unknown length
     9862:  contig  of 4818 bp in length
     9863:  gap     of unknown length
    14680:  contig  of unknown length
    14681:  gap     of unknown length
    14780:  contig  of 3730 bp in length
    14781:  gap     of unknown length
    18510:  contig  of unknown length
    18511:  gap     of unknown length
    18610:  contig  of 5396 bp in length
    18611:  gap     of unknown length
    24006:  contig  of unknown length
    24007:  gap     of unknown length
    24106:  contig  of 7822 bp in length
    24107:  gap     of unknown length
    31928:  contig  of unknown length
    31929:  gap     of unknown length
    32028:  contig  of 13611 bp in length
    32029:  gap     of unknown length
    45639:  contig  of unknown length
    45640:  gap     of unknown length
    45739:  contig  of 9073 bp in length
    45740:  gap     of unknown length
    54812:  contig  of unknown length
    54813:  gap     of unknown length
    54912:  contig  of 11281 bp in length
    54913:  gap     of unknown length
    66193:  contig  of unknown length
    66194:  gap     of unknown length
    66293:  contig  of 16904 bp in length
    66294:  gap     of unknown length
    83197:  contig  of unknown length
    83198:  gap     of unknown length
    83297:  contig  of 23503 bp in length
    83298:  gap     of unknown length
   106800:  contig  of unknown length
   106801:  gap     of unknown length
   106900:  contig  of 38661 bp in length
   106901:  gap     of unknown length
   145561:  contig  of unknown length
   145562:  gap     of unknown length
   145661:  contig  of 44549 bp in length
   145662:  gap     of unknown length
   190210:  contig  of unknown length
```

```
     misc_feature    1130..3107
                     /note="assembly_name:Contig15"
     misc_feature    3208..4755
                     /note="assembly_name:Contig22"
     misc_feature    4856..7087
                     /note="assembly_name:Contig18"
     misc_feature    7188..9762
                     /note="assembly_name:Contig19"
     misc_feature    9863..14680
                     /note="assembly_name:Contig20"
     misc_feature    32029..45639
                     /note="assembly_name:Contig24"
     misc_feature    14781..18510
                     /note="assembly_name:Contig21"
     misc_feature    45740..54812
                     /note="assembly_name:Contig25"
     misc_feature    18611..24006
                     /note="assembly_name:Contig22"
     misc_feature    54913..66193
                     /note="assembly_name:Contig26"
     misc_feature    66294..83197
                     /note="assembly_name:Contig27"
     misc_feature    24107..31928
                     /note="assembly_name:Contig23"
     misc_feature    83298..106800
                     /note="assembly_name:Contig28"
     misc_feature    106901..145561
                     /note="assembly_name:Contig29"
     misc_feature    106901..145561
                     /note="assembly_name:Contig30"
     misc_feature    145662..190210
                     /note="assembly_name:Contig31"
ORIGIN
```

```
RESULT 8
LOCUS       AF361354     1386 bp    mRNA    linear   PRI 21-DEC-2001
DEFINITION  Homo sapiens voltage-dependent calcium channel gamma-8 subunit
            (CACNG8) mRNA, complete cds.
ACCESSION   AF361354
VERSION     AF361354.1  GI:17974541
```

```
Query Match      36.0%;  Score 236;  DB 2;  Length 190210;
Best Local Similarity   79.0%;
Matches  319;  Conservative   0;  Mismatches  80;  Indels  5;  Gaps  3;

Qy         73 GCCCCCCAGCTTCTGCCTGTGCTGTGATCCCCCCAGCGCGCGGACGGGCCCCGCCTTCCGC 132
              ||||||||||||||||||||| |||||||||||||||||||||||||| ||||||| ||||
Db     145831 GGCACCCCGCCTTCTCGGCTGCGCGCTTAAAAAACCCCCCAGCGCCACGCCCCGCCCCCGC 145890

Qy        133 TGCCCCGGTGGTGGCCCCACGGGCCCCCCCCGGTTGCCAGTGGTGGTGAGACTGGAGTCATTGAAAC 192
              ||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||| |||||
Db     145891 TGCCCCGGATGATGCCCACGGGCCCCCCCCGGGCTGCCCCGTGGTCAAAACTGGGAGTGCGTGAAAGC 145950

Qy        193 GCTGGAATCAGAGGAGAGGGGGTTTGTGGGTGTGAAAAAGGGCGCCTTCAGGTACTACTGACCACCA 252
              ||||||||||||||||||||||||||| || |||||||| ||||||||||||||||||||||||||||
Db     145951 GCTGGAACGAAGAGCGGGGGCGCCTCTCGGTCGCGAAGGCGGTGCAGGTGCTCCTGAGGACGG 146010

Qy        253 TCGGGCGCCTTCGCGCGCGCTTTTGGCCTTCATGACCATGGCCATCAGACTGACTACTGGCTCT 312
              |||||||| |||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146011 TGAGGCGCCTTCGC--CTTGTCGGCGCCTCATGACCATCAGCACTGACTACTACTGGCTCT 146068

Qy        313 ACACAAGAGCTTCTCATCTCTGCAACACCACCACCA--CCTCACAGCAGTGATGACGGACCACC 370
              |||||||||||||||||||||||||||||||||||||    |||||||||| |||||||||||||
Db     146069 ACACGCGCCGCCCTCATCTCGCAACACCACACGGGCCGCGCGCGCGCGGGGACCCC 146128

Qy        371 CCATCGTGGGGGGCGCAGTGGCCTCTCCGGAGGAAGGACGCGGGCGGCCCTCACACATTCAGG 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146129 CCACCGCGGGGGGC--GCGGGCGCCTCGGGAGAAGAAGGACCCCGGGGGCGGCCCTCACGCACTCGGG 146187

Qy        431 CCTCTGGCCGGATATGCTGCCTGGAAGGGTAGGGGTGCAGGAGAGGCC 474
              ||||||| |||||||| ||| |||||||| |||| ||| || ||||||
Db     146188 CCTCCTGGAGGATCTGCCTGGAAGGTAGGGGTGCGGGGCGGGGGCCC 146231
```

```
KEYWORDS       .
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 1386)
  AUTHORS      Chu,P.J., Robertson,H.M. and Best,P.M.
  TITLE        Calcium channel gamma subunits provide insights into the evolution
               of this gene family
  JOURNAL      Gene 280 (1-2), 37-48 (2001)
  MEDLINE      21601102
   PUBMED      11738816
REFERENCE    2 (bases 1 to 1386)
  AUTHORS      Chu,P.-J., Robertson,H. and Best,P.M.
  TITLE        Direct Submission
  JOURNAL      Submitted (14-MAR-2001) Department of Molecular and Integrative
               Physiology, University of Illinois, 407 S. Goodwin Ave., Urbana, IL
               61801, USA
FEATURES             Location/Qualifiers
     source          1..1386
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1..1386
                     /gene="CACNG8"
     CDS             106..1386
                     /gene="CACNG8"
                     /codon_start=1
                     /product="voltage-dependent calcium channel gamma-8
                     subunit"
                     /protein_id="AAL50049.1"
                     /db_xref="GI:17974542"
                     /translation="MESLKRWNEERGLWCEKGVQVLLTTVGAFAAFGLMTIAISTDYW
                     LYTRALICNTLNTAGGDDTPPPAGAGGASEKKDPGGLTHSGLWRICCLEGLKRGVC
                     VKINHFPEDTDYDHDSAEYLLRVVRASSIFPILSAILLLLGGVCVAASRVYKSKRNII
                     LGAGILFVAAGLSNIIGVIVYISANAGEPGPKPDEEKKONHYSYGWSFYFGGLSFILAE
                     VIGVLAVNIYIERSREAHCQSRSDLLKAGGGAGGSPSAILRLPSYRFRYRRRSR
                     SSSRSSEPSPSRDASPGGPGPGFASTDISMYTLSRDPSKGSVAAGLAGDAGGGGGGAV
                     GAFGGAAGGAGGGGGGEGAERDRGGSSGFLILHNAFPKEAGGGVTVTVTRPPAPP
                     APRHPAPSAPAPGTLAKGAAASNTNLNRKTTPV"
ORIGIN
Query Match       35.4%;  Score 232.2;  DB 9;  Length 1386;
Best Local Similarity   78.3%;
Matches  306;  Conservative    0;  Mismatches   78;  Indels    7;  Gaps    2;

Qy         73 GCCCCCCAGCTTCTGCCTGTGCTTGCTGATCCCCCCCAGCGCGCACGGCGCCCGCCTCGC 132
              ||||||| || | ||| |||||||||    | || ||||||||| |||| ||||||| ||
Db          3 GGCACCCCCGCTTCTTGGCTGCCGCTTAAAAACCCCCCAAGCGCGCACGCCCCGCCCCGC 62

Qy        133 TGCCCCCGGTGTGGGGCCCACGCGCCCCCCCCCGGTTGCCCAGTGGTGGAAACTGGCGCG 192
              ||||| || |  ||  ||| ||||||||||  ||||| |||||||||||| |||   |||
Db         63 TGCCCCGGATGATG--CCACGGCCCCCGTGGTCCGCTGCCTCGCCCTGTCAAACTGGGAGC 121

Qy        193 GCTGGAATGAAGAGGAGGGGGTTGTGGTGGTGAAAAGGGGTTCAGGTACTACTGACACCA 252
              |||||||   || | || |||||||||||||| |||||||||||||||||||||||||||
Db        122 GCTGGAACGAAGAGCGGGGGCCTCTGGTGTGCGAGAAGGGGGTTCAGGTGCTGCTTGACGACGG 181

Qy        253 TCGGCGCCTTCCGCGGCTTTTGGCCTCATGACCACTGACTACTACTGGCTCT 312
              | |||||||||||| |||| ||||| || || ||| |||||||||||||||||
Db        182 TGGGCGCCTTCGGCCGCCTTCATGATCACTGCCATCAGCACTGACTACTGGCTCT 241

Qy        313 ACACAAGAGAGCTCTCATCTGCAACCACCACCTCACAGACAGGTGATGACG-------GAC 366
              || || |||| |  | || | |||||| ||||| || || ||||||||       |||
Db        242 ACACGCGCGCCCTCATCTGCAACACCCTCAACACCGCCGGGGGGGATGACACCCCGCCGCCC 301

Qy        367 CACCCCATCGTGGGGGACGAGAGTGGGCTGGGGCTGGGGGCTCTCCGAGAGAAGAAGGACC 426
              |||||| || || |||||| ||||||||| |||| |||||||||||||||| || |||||
Db        302 CTCCACCTCGGCGGCTGGCGCTGGCGCTGGGGGCTCTCCGAGAGAAGAAGGACCCCGGCGG 361

Qy        427 CAGGCCCTCTGGCCCGGATATGCCTGCCTGGAAGG 457
              |||||| | ||  |||| |||| |||||| |||||
Db        362 CGGGCCCTCTGAGGGGATCTGCCTGCCTGGAAGG 392
```

```
RESULT 9
LOCUS       AF234892                 633 bp    mRNA    linear    PRI 02-JAN-2002
DEFINITION  Homo sapiens putative voltage gated calcium channel gamma-8 subunit
            (CACNG8) mRNA, partial cds.
ACCESSION   AF234892
VERSION     AF234892.1  GI:13182981
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 633)
  AUTHORS   Black,J.L. III, Kryzer,T.J. and Lennon,V.A.
  TITLE     Proposed Homo sapiens voltage-gated calcium channel gamma-6 subunit
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 633)
  AUTHORS   Black,J.L. III, Kryzer,T.J. and Lennon,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Department of Psychiatry and Psychology,
            Mayo Clinic, 200 SW 1st Street, Rochester, MN 55905, USA
FEATURES             Location/Qualifiers
     source          1..633
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="cerebellum"
     CDS             <1..>633
                     /codon_start=1
                     /product="putative voltage gated calcium channel gamma-8
                     subunit CACNG8"
                     /protein_id="AAK15019.1"
                     /db_xref="GI:13182982"
                     /translation="GLNRPPGQVSLKRWNEERGLWCKKGVQVLLTTVGAAFGLMTI
                     AISTDYWLYTRALICNTNTITAGGDGTPHRGGAGGASEKKDPGGLTHSGLWRICCLEG
                     LKRGVCVKINHFPEDYDHDSAEYLLRVVRASSIFPILSAILLLLGGVCVAASRVYK
                     SKRNIILGAGILFVAAGLSNIIGVIVYISANAGEPGPKPDEEKKONHYSYGW"
     misc_feature    76
                     /note="possible alternate start site"
ORIGIN
Query Match       29.1%;  Score 191.2;  DB 9;  Length 633;
Best Local Similarity   82.0%;
Matches  233;  Conservative    0;  Mismatches   48;  Indels    3;  Gaps    1;

Qy        177 CTGGAGTCATTGAAACGCTGGAATGAAGAGGAGGGGGTTGTGGTGGTGAAAAGGGGTTCAG 236
              || |||||||||||||| ||||||||| |||||||||||||||||||||||   |||||||
Db         22 CAGGTGTCGCTGAAGCGCTGGAACGAAGAGCGGGGGCTTTGCTCATGACCATCAGC 81

Qy        237 GTPACTACTGACCACCATCGGCGCCTTCGCGGTGTTTTGGCCTCATGACCATCAGC 296
              |||||||||||| |||| || |||| | | ||| | |||||| || ||||||||||
Db         82 GTGCTGCTTGACGACGGTGGGCGCCTTCGCCGCCTTCATGACCATCAGC 141

Qy        297 ACTGACTACTGGCTCTCTACACAAGAGAGCTCTCATCTGCAACACCACCTCACAGC----A 353
              ||||||||||||||||||||| || | || |  | || | |||||| || ||||    |
Db        142 ACTGACTACTGGCTCTACACGCGCGCCCTCATCTGCAACACCCTCAACACCGCCGGGGGGG 201

Qy        354 GGTGATGACGGACGACCACCCCATCGTGGGGGACGAGAGTGGGCTGGGGCTGGGG 413
              || || |||| ||| || |||||| || |||||| ||||||||| |||| ||||
Db        202 GCGGACGACGGACAGCCCCCACCGGCGGCTGGCGCTGGCGCTGGGGGGCTGGGG 261

Qy        414 GGCCTCACACATTCAGGCCCTCTGGCCCGGATATGCCTGGAAGG 457
              || |  | ||| || |||||| |||| |||||||| |||| ||||
Db        262 GGCCTCACGGCACTCGGGGCCCTCTGGAGATCTGCCTGGAAGG 305
```

```
RESULT 10
LOCUS       AF288388                1281 bp    mRNA    linear    PRI 16-MAR-2001
DEFINITION  Homo sapiens calcium channel gamma subunit 8 (CACNG8) mRNA, partial
            cds.
```

ACCESSION   AF288388
VERSION   AF288388.1 GI:13357179
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1281)
  AUTHORS   Burgess,D.L., Gefrides,L.A., Foreman,P.J. and Noebels,J.L.
  TITLE   A cluster of three novel Ca2+ channel gamma subunit genes on
            chromosome 19q13.4: evolution and expression profile of the gamma
            subunit gene family
  JOURNAL   Genomics 71 (3), 339-350 (2001)
  MEDLINE   21100909
   PUBMED   11170751
REFERENCE   2 (bases 1 to 1281)
  AUTHORS   Burgess,D.L., Gefrides,L.A., Foreman,P.J. and Noebels,J.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-JUL-2000) Neurology, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
FEATURES   Location/Qualifiers
     source   1..1281
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /map="19q13.4"
     gene   1..1281
                     /gene="CACNG8"
     CDS   <1..1245
                     /gene="CACNG8"
                     /codon_start=1
                     /product="calcium channel gamma subunit 8"
                     /protein_id="AAK20031.1"
                     /db_xref="GI:13357180"
                     /translation="GLWCEKGVQVLLTVGAFAAFGLMTIAISTDYWLYTRALICNTT
                     NLTAGDDGTPHRGGGGASEKDPGGLTHSGLWRICCLEGLKRGVCVKINHFPEDTDY
                     DHDSAEYLLRVVRAASSFPILSAILLLGGVCVAASRVYKSKRNILGAGILFVAAGL
                     SNIIGVIYISANAGEPGPKRDEEKKNHYSYGWSFYFGGLSFILAEVIGVLAVNIYIE
                     RSREAHCQSRDLLKAGGGAGGSSGGSGPSAILRLPSYRFRYRRRSRSSSRSSEPSPSR
                     DASPGGPGFRASTDISMYLSRDPSKGSVAAGLAGGGGGGGGGAVGAFGGAAGGAGG
                     GGGGGGGAAERDRGGASGFLTLHNAFPKEAGGGVTVTVTGPPAPAPAPAPSAPAP
                     GTLAKEAAASNTNTLNRKTTPV"
ORIGIN

Query Match   26.1%; Score 171; DB 9; Length 1281;
  Best Local Similarity   82.9%; Pred. No. 8.1e-30;
  Matches 208; Conservative 0; Mismatches 40; Indels 3; Gaps 1;

Qy   210 GGTTTGTGGTGTGAAAAGGGCGTTCAGTTACTACTGACCATCATCGGGCGCCCTTGGCGGGCT 269
          ||||| || || || ||||||||||||||||||||||||||| |||||||||| ||| |
Db   1 GGCCTCTCGTCGCGAGAGGGGGGTGCAGGGTGCTCAGTTACTACGCGTGGGCGCCCTTCGCCGCCC 60

Qy   270 TTTTGGCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGCTTCTCATC 329
          | ||||||||||||||||||||||||||||||| ||||||||||||||| || ||| || | |
Db   61 TTCGGCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACACGCGCGGCGCCCTCATC 120

Qy   330 TGCAACACACACACCAACCCTCACAGC---AGGTGATGACGGACCACCACCCCATCGGTGGGGGCAGT 386
          ||||||||||||| || || |||| || | ||| |||||||| || ||||||||| ||||||||
Db   121 TGCAACACACACCTCACGGCCCGCGGCCGGACGCGACGGGACCACCACCCCATCGGTGGGGGCGGC 180

Qy   387 GGCTTCCTCCGAGAAGAAGGACCCTGGGGGGCTCACGATTCAGGGCCTGGGCGGGATATGC 446
          || | ||| ||||||||||||||| |||| ||||||| || |||||||||||||||| ||||
Db   181 GGGGCCCTCGGAGAAGAAGGACACCCGGGCCCTCACGCACTCGGGCCTGGGCGGGATCTGC 240

Qy   a 447 TGCCTGGGAAGG 457
          ||||||||||||
Db   241 TGCCTGGGAAGG 251

RESULT 11
AY037891
LOCUS   AY037891   987 bp   mRNA   linear   VRT 01-APR-2002

DEFINITION Gallus gallus calcium channel gamma 4 subunit (CACNG4) mRNA,
            complete cds.
ACCESSION   AY037891
VERSION   AY037891.1 GI:15418940
KEYWORDS   .
SOURCE   Gallus gallus (chicken)
  ORGANISM Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 987)
  AUTHORS   Kious,B.M., Baker,C.V., Bronner-Fraser,M. and Knecht,A.K.
  TITLE   Identification and characterization of a calcium channel gamma
            subunit expressed in differentiating neurons and myoblasts
  JOURNAL   Dev. Biol. 243 (2), 249-259 (2002)
  MEDLINE   21881550
   PUBMED   11884034
REFERENCE   2 (bases 1 to 987)
  AUTHORS   Kious,B.M., Bronner-Fraser,M. and Knecht,A.K.
  TITLE   Direct Submission
  JOURNAL   Submitted (04-JUN-2001) Biology, Caltech, Mail Code 139-74,
            Pasadena, CA 91125, USA
FEATURES   Location/Qualifiers
     source   1..987
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
     gene   1..987
                     /gene="CACNG4"
     CDS   1..987
                     /gene="CACNG4"
                     /codon_start=1
                     /product="calcium channel gamma 4 subunit"
                     /protein_id="AAK72510.1"
                     /db_xref="GI:15418941"
                     /translation="MVWCDRGVQMLLTTVGAFAAFSLMAIAIGTDYWLYSSAHICNGT
                     NTTTDTDQGPPKKARGDLTHSGLWRICCLEGLYKGHCCRINHFPEDNDYDHDSSEYLL
                     RIVRAASSVFPILSTILLLGLCVGAGRIYNSKKNIILSAGILFVAAGLSNIIGIIVY
                     ISSNAGDPSPSDKRDEDKKNHYNYGWSFYFGALSFIVAETIGVLAVNIYIEKNKELRFKT
                     KREFLKTSSSSPYARMPSYRYRRRSRSSSRSTEPSPSRDISPVGMKIASTIPMNEIS
                     MYTLSREPLKVTTAASYNADQEASFLQVENFLQKEFKEGLHVMNVRRTTPV"
ORIGIN

Query Match   14.0%; Score 92; DB 5; Length 987;
  Best Local Similarity   65.9%; Pred. No. 3.7e-11;
  Matches 162; Conservative 0; Mismatches 60; Indels 24; Gaps 1;

Qy   214 TGTGGTGTGAAAAGGGCGTTCAGTTACTACTGACCATCATCGGGCGCCCTTGGCGGGCTTTTG 273
          |||||||||| | || ||||| |||| |||||||| ||| |||||||||||| |||| |
Db   5 TGTGGTGTGACCGTGGCGTCCAGATGCTGCTGACCACGGTGGGGAGCTTTGGCGGGGCTTTCA 64

Qy   274 GCCTCCATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGCTTCTCATCTGCA 333
          |||||||||| ||||||||| | |||||||||||||||||| | || || || |||||||
Db   65 GCCTCCATGGCCATCGCCATCGGTACGGACTACTGGCTCTACTCCAACGGCGACATCGTATGCA 124

Qy   334 ACACCACCAACCTCACAGCAGGTGATGACGGACCACCACCCCATCGGTGGGGGCAGTGGCTCCT 393
          |||||||||||||||||| | ||| |||||||||||||| | ||||||||| |||| |
Db   125 ACGGCACCAACATCACGACAGCGGGACGATACCCAGGGGCCGC-----------------163

Qy   394 CCGAGAAGAAGGACCCTGGGGGGCTCACGATTCAGGGCCTGGGCGGGATATGCTGCCTGG 453
          |||||| || |||| | |||| ||||| ||| |||||||||||||||| | |||||||
Db   164 ---CGAAGAAAGCGAGGGGCGGGCGACTCACGCATTCGGGGGCTTGGGGACTGCTGCCTCG 220

Qy   454 AAGGTA 459
          |||| |
Db   221 AAGGAA 226

RESULT 12
LOCUS   HSCACNG1
HSCACNG1   720 bp   DNA   linear   PRI 07-JAN-2000
DEFINITION Homo sapiens calcium channel gamma 4 subunit (CACNG4) gene, exon 1.
ACCESSION   AF142622

VERSION   AF142622.1 GI:6062998
KEYWORDS   .
SEGMENT   1 of 4
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 720)
  AUTHORS   Burgess,D.L., Davis,C.F., Gefrides,L.A. and Noebels,J.L.
  TITLE   Identification of three novel Ca(2+) channel gamma subunit genes
             reveals molecular diversification by tandem and chromosome
             duplication
  JOURNAL   Genome Res. 9 (12), 1204-1213 (1999)
  MEDLINE   20082967
   PUBMED   10613843
REFERENCE   2 (bases 1 to 720)
  AUTHORS   Burgess,D.L., Caleb,D.F., Lisa,G.A. and Jeffrey,N.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-APR-1999) Neurology, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
FEATURES   Location/Qualifiers
     source   1..720
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17q22-q24"
     exon   <235..454
                     /gene="CACNG4"
                     /number=1
ORIGIN

Query Match   13.2%; Score 86.4; DB 9; Length 720;
Best Local Similarity   66.8%; Pred. No. 7.8e-10;
Matches 123; Conservative   0; Mismatches 61; Indels   0; Gaps   0;

Qy   214 TGTGGTGTGAAAAGGGCGGTTCAGGTACTACTGACCACCATCGCGCGCCTTCGCGCGCTTTTG 273
          |||| ||||| ||| |||||||||| || ||||||||||||||||||||||||| |||
Db   239 TGCGGATGCGACCGCGCGGCTGCAGATGCTGCCAGATGCTGACCACGCGCCGGAGCCTTCGCGCCCTTCT 298

Qy   274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGCTCTCATCTGCA 333
          |||||||| |||||||||||| || |||||||||||| ||| ||| |||||||||||
Db   299 CGCTTCATGGCCATCGCCATCGGCGACCGACTACTGGCGTGTACTCCAGCGCCGCACATCTGCA 358

Qy   334 ACACCACCAAACCTTCACAGCAGGTGATGACGGACCACCCCATCGTGGGGGGGCAGTGGCTCCT 393
          || |||||||||| ||| | | |||| ||| ||| |||||||| |||||| |||||
Db   359 ACGGCACCAAACCTGACCATGGACGACGCGGGCCCCCGCGCCCCCGCGCGCGGCGGACCTCA 418

Qy   394 CCGA 397
          || |
Db   419 CCCA 422

RESULT 13
LOCUS   AR165149   984 bp   DNA   linear   PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274380.
ACCESSION   AR165149
VERSION   AR165149.1 GI:16238585
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 984)
  AUTHORS   Duckworth,D.Malcolm. and Hayes,P.David.
  TITLE   Cacnglike3 polynucleotides and expression systems
  JOURNAL   Patent: US 6274380-A 1 14-AUG-2001;
FEATURES   Location/Qualifiers
     source   1..984
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   13.2%; Score 86.4; DB 6; Length 984;
Best Local Similarity   66.8%; Pred. No. 7.8e-10;
Matches 123; Conservative   0; Mismatches 61; Indels   0; Gaps   0;

Qy   214 TGTGGTGTGAAAAGGGCGGTTCAGGTACTACTGACCACCATCGCGCGCCTTCGCGCGCTTTTG 273
          |||| ||||| ||| |||||||||| || ||||||||||||||||||||||||| |||
Db   5 TGCGGATGCGACCGCGCGGCTGCAGATGCTGCCAGATGCTGACCACGCGCCGGAGCCTTCGCGCCCTTCT 64

Qy   274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGCTCTCATCTGCA 333
          |||||||| |||||||||||| || |||||||||||| ||| ||| |||||||||||
Db   65 CGCTTCATGGCCATCGCCATCGGCGACCGACTACTGGCGTGTACTCCAGCGCCGCACATCTGCA 124

Qy   334 ACACCACCAAACCTTCACAGCAGGTGATGACGGACCACCCCATCGTGGGGGGGCAGTGGCTCCT 393
          || |||||||||| ||| | | |||| ||| ||| |||||||| |||||| |||||
Db   125 ACGGCACCAAACCTGACCATGGACGACGCGGGCCCCCGCGCCCCCGCGCGCGGCGGACCTCA 184

Qy   394 CCGA 397
          || |
Db   185 CCCA 188

RESULT 14
LOCUS   AX101260   984 bp   DNA   linear   PAT 10-APR-2001
DEFINITION Sequence 1 from Patent WO0121791.
ACCESSION   AX101260
VERSION   AX101260.1 GI:13620050
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Clare,J.J., Plumpton,M., Moss,F.J. and Sanseau,P.
  TITLE   Stargazin-like neuronal ca2+-channel gamma subunit polypeptides
  JOURNAL   Patent: WO 0121791-A 1 29-MAR-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES   Location/Qualifiers
     source   1..984
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS   1..984
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC36506.1"
                     /db_xref="GI:13620051"
                     /db_xref="REMTREMBL:CAC36506"
                     /translation="MVRCRDRGLQMLTTAGAFAAFSLMAIAIGTDVWLYSSAHICNGT
                     NLTMDDGPPRRARGDLTHSGLWRVCCXXGIYKGHCFRINIFPEDNDYDHDSSEYLLR
                     IVRASSVPFILSTILLLLGCLCIGAGRIYSRKNNIVLSAGILFVAAGLSNIIGIIVYI
                     SSNTGDPSDKRDEDKKQNHYNYGWSFYFGALSFIVAETVGVLAVNIYIEKNKELRFKTK
                     REFLKASSSSPYARMPSYRYRRRSRSSSRSTEASPSRDVSPMGLKITGAIPMGELSM
                     YTLSREPLKVTTAASYSPDQEASFLQVHDFPQQDLKEGFHVSMLNRRTTPV"
ORIGIN

Query Match   13.2%; Score 86.4; DB 6; Length 984;
Best Local Similarity   66.8%; Pred. No. 7.8e-10;
Matches 123; Conservative   0; Mismatches 61; Indels   0; Gaps   0;

Qy   214 TGTGGTGTGAAAAGGGCGGTTCAGGTACTACTGACCACCATCGCGCGCCTTCGCGCGCTTTTG 273
          |||| ||||| ||| |||||||||| || ||||||||||||||||||||||||| |||
Db   5 TGCGGATGCGACCGCGCGGCTGCAGATGCTGCCAGATGCTGACCACGCGCCGGAGCCTTCGCGCCCTTCT 64

Qy   274 GCCTCATGACCATCGCCATCAGCACTGACTACTGGCTCTACACAAGAGCTCTCATCTGCA 333
          |||||||| |||||||||||| || |||||||||||| ||| ||| |||||||||||
Db   65 CGCTTCATGGCCATCGCCATCGGCGACCGACTACTGGCGTGTACTCCAGCGCCGCACATCTGCA 124

Qy   334 ACACCACCAAACCTTCACAGCAGGTGATGACGGACCACCCCATCGTGGGGGGGCAGTGGCTCCT 393
          || |||||||||| ||| | | |||| ||| ||| |||||||| |||||| |||||
Db   125 ACGGCACCAAACCTGACCATGGACGACGCGGGCCCCCGCGCCCCCGCGCGCGGCGGACCTCA 184

Qy   394 CCGA 397

RESULT 15

```
Db          |||
        185 CCCA 188
```

```
LOCUS           MMU272045                984 bp    mRNA    linear   ROD 04-APR-2000
DEFINITION      Mus musculus mRNA for calcium channel gamma 4 subunit (CACNG4
                gene).
ACCESSION       AJ272045
VERSION         AJ272045.1  GI:7452995
KEYWORDS        CACNG4 gene; calcium channel gamma 4 subunit.
SOURCE          Mus musculus (house mouse)
  ORGANISM      Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1  (bases 1 to 984)
  AUTHORS       Klugbauer,N., Dai,S., Specht,V., Lacinova,L., Marais,E., Bohn,G.
                and Hofmann,F.
  TITLE         A family of gamma-like calcium channel subunits
  JOURNAL       FEBS Lett. 470 (2), 189-197 (2000)
   MEDLINE      20020013
   PUBMED       10734232
REFERENCE       2  (bases 1 to 984)
  AUTHORS       Klugbauer,N.
  TITLE         Direct Submission
  JOURNAL       Submitted (08-FEB-2000) Klugbauer N., Institut fuer Pharmakologie
                und Toxikologie, Technische Universitaet Muenchen, Biedersteiner
                Str. 29, 81827 Muenchen, GERMANY
FEATURES             Location/Qualifiers
     source          1..984
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="brain"
     gene            1..984
                     /gene="CACNG4"
     CDS             1..984
                     /gene="CACNG4"
                     /codon_start=1
                     /product="calcium channel gamma 4 subunit"
                     /protein_id="CAB86386.1"
                     /db_xref="GI:7452996"
                     /db_xref="GOA:Q9JJV4"
                     /db_xref="SWISS-PROT:Q9JJV4"
                     /translation="MVRCDRGLQMLLTTPAGAFAAFSLMAIAIGTDYWLYSSAHICNGT
                     NLTMDGPPPRARGDLTHSGLMRVCCIEGIYRGHCRRINHFPEDNDYDHDSSEYLLR
                     IVRASSVPFILSTILLLLGGLCIGAGRIYSRKNIVLSAGILFVAAGLSNIIGIIVYI
                     SNTGDPSDKRDEDKKQHYNYGWSFYGALSFIVAETVGVLAVNIYIEKNKELRFKTK
                     REFLKASSSSPYARMPSYRYRRRSRSSRSTEASPSRDASPVGLKITGAIPMGELSM
                     YTLSREPLKVTTPAASYSPDQDAGFLQMHDFFQQDLKEGFHVSMLNRRTTPV"
ORIGIN
```

```
Query Match               13.2%;   Score 86.4;  DB 10;   Length 984;
Best Local Similarity     66.8%;   Pred. No. 7.8e-10;
Matches  123;  Conservative   0;  Mismatches  61;  Indels   0;  Gaps   0;
```

```
Qy    214  TGTGGTGTGAAAAGGGCGGTTCAGTTACTTACTGACCACCATCGGGGCCCTTCGCGCGTTTG  273
           || ||||||| ||| |||||||   |||| ||||| |||||||||||||   |||||||||
Db      5  TGCGGATGCGACCGCGCGGCCTGCACGATGCTGCTGACCACGGACGCCTTCGCCCGCTTCT   64

Qy    274  GCCTCCATGACCATCCCCATCAGCACTGACTGGCCTCTACACAAGAGCTTCATCTGCA  333
           ||||||||||||| ||||||||||| |||||||||||||| |||||| || | |||||
Db     65  CGCTCCATGGCCATCGCCATCGGCACCGACTACTGGCTGTACTCCAGCGCGCACATCTGCA  124

Qy    334  ACACCACCAACCTCACAGCAGGTGATGACGGACCACCCCATCGTGGGGCCAGTGGCTCCT  393
              |||||||||||| ||| || | |||||| || |||| || |||| | |||| ||||
Db    125  ACGGCACCAACCTGACCATGGACGGACCGCCCCCCCGCGCCCGCGGCGACCTCA  184

Qy    394  CCGA  397
           || ||
Db    185  CCCA  188
```